

Suggestions for
THE USE OF
NEW-TYPE TESTS
in India

BY

EMIL W. MENZEL

B.D., M.A. IN EDUCATION



HUMPHREY MILFORD
OXFORD UNIVERSITY PRESS

1942

OXFORD UNIVERSITY PRESS
AMEN HOUSE, LONDON, E. C. 4
EDINBURGH GLASGOW NEW YORK
TORONTO MELBOURNE CAPE TOWN
BOMBAY CALCUTTA MADRAS
HUMPHREY MILFORD
PUBLISHER TO THE
UNIVERSITY

First published 1930
Second edition 1942

TO
MANY SMILING EXAMINEES

P R E F A C E

EXAMINATIONS in India are undoubtedly in need of drastic reform. On 28 July 1937, *The Statesman* in Calcutta carried the headlines :

NEW METHOD OF RECRUITMENT TO GOVERNMENT SERVICES

INTELLIGENCE TEST AT AGE OF 17

Revolutionary Proposal

About the same time that the above headlines appeared, the Wardha Scheme of education was published. The two following paragraphs, quoted from this Scheme, do not stand alone in condemning the present examination system and proposing a more objective method of examining but express what educationists all over India have felt to be true, although few may want to say this as strongly as the Wardha Scheme does.

The system of examination prevailing in our country has proved a curse to education. A bad system of education has, if possible, been made worse, by awarding to examinations a place out of all proportion to their utility. As a measure of the work of individual pupils or the schools, by a consensus of expert opinion, examinations are neither valid nor complete. They are inadequate and unreliable, capricious and arbitrary. We shall take care to guard the proposed system of general national education against their baneful influence.

The purpose of the examination can be served by an administrative check of the work of the schools in a prescribed area, by a sample measurement of the attainment of selected groups of students conducted by the

inspectors of the Education Board. The tests so administered should be constructed in close consultation with the specialists responsible for curriculum revision. They should be long enough to cover the whole range of the curriculum and should be in a form which makes marking objective and independent of individual judgement.¹

Regardless of one's attitude towards the Wardha Scheme as a whole, it must be admitted that the Scheme has very good reasons for claiming that there are truly baneful influences of the present examination system

But the effects of examinations need not be thus baneful. Education will not get along without examinations of some sort or other. But these can be greatly improved and there is a well-tried technique already available which can greatly improve the examination system provided it is used with discretion

An improvement of the examination system is only a part of the benefit to be derived from the use of New-Type tests. Of fully as great benefit as the improvement of the examination system is the improved method of measuring educational progress or retrogression and psychological reaction. Entire new fields for educational effort have been opened up by an improved method of measuring educational effects. Testing is not only examining for the sake of deciding which pupils shall 'pass' and which shall 'fail', but can be made a technique of measurement which enables one to see what is happening in the minds, and (to some extent) the hearts of the examinees.

Modern education cannot claim to have a single important idea that was not known and used several generations ago. Modern education can only claim to have further developed a few old ideas and to have improved

¹ *Harijan*, 11 December 1937.

the technique of measuring mental and educational changes to the point where the measurement can be said to be relatively scientific. The considerable progress which has taken place during the last generation in education is largely due to the fact that we no longer need to guess at the extent of mental and educational effects but can actually measure many of them. The improved way of measuring has made the difference between mere debate and knowledge, between guess-work and a reasonable certainty as to the soundness or unsoundness of many of our educational practices.

India can hope to make little progress towards modernizing her educational practices until she too improves her way of measuring. There may be some who will say that India does not want modern education in the sense of imported practices from the West, but that she wishes to develop an indigenous culture and an indigenous system of education. No true educator would withhold from this country her right to seek an educational system better adapted to her peculiar problems than the present system. It is precisely because this country does need to find a more naturally Indian system of education that improved devices of measurement are needed so that the experimenter can recognize the educational pitfalls and blind alleys which fritter away the lives of teachers and pupils. Whatever direction education in India may pursue, it will need the help of scientific measurement in order to progress towards a worthy goal. It is the pioneer and experimenter who need to be able to measure their progress or retrogression, not the blind imitator and mere preserver of the *status quo*. It is the explorer who uses compass, sextant, and log-book, not the 'stay-at-home'.

There are many excellent references in the English language on 'How to Measure in Education'. Hardly a

month goes by without a capable reference being added to the list. What is then the purpose of adding another volume to the shelf?

Among all the excellent references to be had there is not one that takes cognizance of the Indian situation. It is one thing to understand the results of an Intelligence test as given in England and America and another thing to understand what has happened when a similar test is given in India to boys and girls whose cultural, social, and physical background is entirely different. It is one thing to judge school accomplishment on the part of pupils in a Northumberland or New York school and quite another to judge in an ordinary rural school in India. Whatever is used in India must be thoroughly adapted. There are a thousand and one questions of adaptation which come up when we transplant the measuring devices of other countries to Indian soil. It is to call attention to some of these that the present book has been written and also to put into the hands of the progressive Indian teacher a brief guide that does not assume on the part of the reader the familiarity with psychological literature and modern methods of measurements in education which the English or American author assumes in the audience he is writing for.

There is need for a guide-book to New-type testing and measurement which keeps in mind the virgin soil which constitutes the Indian problem. What this volume loses in the inexperience of the author in the general field will, it is hoped, be made up by the author's familiarity with Indian educational problems and his experience in having standardized several tests for India.

The author has listed and briefly discussed all the standardizations of tests made in India which have come to his attention up to the date of writing. Undoubtedly

there are some significant experiments about which he does not know. Considerable correspondence was conducted to locate as many of these as possible but it is not easy to find out what has been going on everywhere in such a large country of many languages where the results are very often not published.

Much of that on which the modern testing movement is working is still in the experimental stage while much is established as being scientifically sound and valid. Attempt has been made in these pages, especially in the last chapter, to point out what is considered by scholars to be fairly well 'proved' and what is still regarded as experimental. An attempt has also been made to point out what types of measurements are most likely to prove feasible and useful in India in the near future.

It is hoped that this book may prove useful not only to individuals but to training schools where improved methods of examining and conducting experiments in education are urgently needed to be taught, but have not been taught, perhaps partly because no suitable guide-book for the Indian student has been available.

Chapters I to V will be found to apply most directly to the problems of the classroom teacher although Chapters VI and VII also cover ground very relevant to his problem. Every up-to-date teacher needs to know something about Intelligence tests even though he does not use them himself. Chapters VIII to XI deal with matter that is of less importance to the classroom teacher though by no means unimportant. These chapters are more important to the principal, the student adviser, and administrator. These chapters deal with matters which are still rather new to scientific treatment and which need further development before tests in these fields come into general use.

The writer is indebted to many of his colleagues for assistance in making this book possible. To Mr L. D'Sylva, former Inspector of Schools of the Chhattisgarh Circle and now Deputy Director of Public Instruction of the Central Provinces, he is obligated for a fellowship which is in large part responsible for the initiation of the author into experimentation in education and psychology. He is indebted to Dr J. Manry, Mr W. M. Ryburn, Dr E. L. King, Rev J. C. Koenig, Miss H. Schaeffer, and to his wife for helpful criticism and aid in the final stages of the preparation of the manuscript. Of no less service have been those colleagues, inspectors, deputy inspectors, and headmasters especially Mr B. Sadhu who have spent hours of their time helping to carry out some of the experiments described in this book. Their willing-co-operation has made possible experiment on a large enough scale to be of at least some significance. Last but not least, the thousands of splendid and laughing young sons and daughters of India who took the tests like the heroes and heroines they are, deserve my gratitude. The writer is guilty of appropriating not less than 60,000 hours of their time, for which either a substantial contribution to Indian education or profuse apologies are due.

E. W. M.

CONTENTS

PART I

TESTS FOR THE SCHOOLROOM

CHAPTER	PAGE
I CONDUCTING NEW-TYPE TESTS FOR CLASS	
USE	3
II WHAT IS A STANDARDIZED TEST? ..	30
III SCHOOL-SUBJECT TESTS (STANDARDIZED) ...	43
IV ADVANTAGES AND LIMITATIONS OF NEW-TYPE EXAMINATIONS	88
V THE USE OF DIAGNOSTIC AND INSTRUCTIONAL TESTS	111

PART II

PSYCHOLOGICAL TESTS

VI WHAT ARE INTELLIGENCE TESTS? ...	129
VII INTELLIGENCE TESTS IN USE IN INDIA ...	154
VIII PROGNOSIS AND APTITUDE TESTS... ..	180
IX MEASURING OTHER QUALITIES	191
X SIMPLE STATISTICAL DEVICES AS A HELP IN INTERPRETING RESULTS	205
XI CONCLUSION	227

APPENDICES

I YEARLY TEST FOR CLASS III OF THE PRIMARY SCHOOL	235
II 'THE MIDDLE SCHOOL ACHIEVEMENT TEST' (IN HINDI)	246
III THE CHATTERJI ARITHMETIC TEST ...	261
IV A 'CHARACTER CUP' FOR CLASS VIII ...	266
V INDIAN CLASS SYSTEMS	268
VI BIBLIOGRAPHY... ..	271

PART I

TESTS FOR THE SCHOOLROOM

CONDUCTING NEW-TYPE TESTS FOR CLASS USE

THE best way to arrive at an understanding of modern testing methods is to construct a few informal tests oneself. This any intelligent teacher who will take a little trouble to inform himself, and who enjoys experimenting in a modest way, can quickly learn to do. He will find plenty of scope for exercising his ingenuity, because making a good test is an art.

One does not need to wait until the end of the school term for an excuse to give informal New-Type tests. Tests of this type can be made excellent teaching devices and come as much under the heading 'lesson preparation' as under that of 'final examinations'. The teacher will find them extremely useful in checking his own work as well as that of his pupils. Any teacher who does not merely 'lecture' and 'talk at' his pupils, but strives to give them certain ideas and to help them to develop certain attitudes, will find informal tests, of the type that pupils consider a recreation and not an ordeal, of inestimable value in letting him know what ground has been adequately covered and what ground is to the pupil still enveloped in fog.

I. THE TRUE-FALSE TEST

Let us assume that a class has just finished a study unit of Akbar's reign. The teacher wants to know whether or not Akbar's statesmanship is duly appreciated and whether the main facts are reasonably clear to the pupils. The True-False test will give a good opportunity to see if the pupils

know the facts well enough to identify them in a mass of assertions some of which are correct and some of which are not. The teacher may hand out cyclostyled or printed copies of the test to the pupils but it is by no means necessary to make such elaborate preparations. The test items may be dictated. In order to give the following test of forty questions, the teacher need only instruct the pupils to take a long narrow sheet of paper and write down the consecutive numbers from 1 to 40. As the teacher reads statement number one the pupil puts a plus sign beside No. 1 on his own sheet if he considers the statement correct, or a zero sign if he considers the statement incorrect. Then item number two is read out, and the pupil puts the appropriate sign beside the corresponding number on his own sheet. The whole test can be given in fifteen minutes. If the teacher permits, a lively discussion is bound to follow the giving of the test, clearing up hazy ideas and false impressions. The class will very likely enjoy itself and the time will be well spent.

Weidemann¹ suggests that the following directions be given with a True-False test:

Directions

About one-half of the following statements are true and about one-half are false. Mark each true statement with a plus sign (or an x) on the dotted line at the left of the statement. Mark each statement that is partly or wholly false with a zero (0) on the dotted line at the left of the statement. Do not mark statements that you do not know. Mark the statements in order.

Do not guess. Guessing reduces your score.

Ask no question.

¹ Weidemann, C. C., *How to Construct the True-False Examination* p. 78. Teachers' College Contributions to Education, No. 225. Teachers' College, Columbia University, New York, 1926.

Of course, other wordings are possible. If the warning against guessing is not given, one may expect more guessing. Sometimes this is even required, the examiner requiring every question to be marked one way or the other. The results of a test given to two different groups are comparable only if the directions were the same, for the wording of the directions may influence the score quite appreciably.

- x 1. Akbar was a descendant of Timur.
- o 2. The father of Akbar was Babur.
- o 3. Akbar spent most of his youth in the palace at Delhi.
- o 4. Akbar was the first man to call himself 'Kaisar-i-Hind'.
- x 5. Akbar was a very young man when he became emperor.
- o 6. Akbar's regent was Sher Shah.
- x 7. Akbar treated the rebelling Bairam kindly.
- x 8. When Akbar ascended the throne his empire was much smaller than that of his grandfather.
- o 9. Akbar's empire stretched to Cape Comorin.
- x 10. Akbar tried to found a new religion.
- o 11. Akbar reserved the highest places in his government for Mohammedans.
- x 12. Much of Akbar's system of government survives until today.
- o 13. Akbar encouraged that part of the Hindu community in which suttee was sometimes observed to continue this practice.
- x 14. Akbar said: 'Why should I claim to guide men before I myself am guided?'
- x 15. Akbar once rode more than 200 miles on horse-and camel-back in two days.
- o 16. Akbar is unique more because of his generalship displayed in war than for his work in peacetime.
- o 17. Akbar was able to consolidate the empire because he treated rebels and enemies with the strongest discipline.

- x 18. Although Akbar was a Mohammedan he put Hindus and Parsees in responsible positions.
- o 19. Akbar was blessed with well-behaved sons.
- o 20. As he was a man of war, Akbar had little time for poets and scholars
- o 21. The most beautiful building that Akbar caused to be built is the Taj Mahal.
- x 22. Though he knew little Sanskrit, Akbar was interested enough in Sanskrit literature to have much of it translated into Persian.
- x 23. Akbar spent some years of his life in exile.
- x 24. His capital city throughout his reign was Delhi.
- o 25. His abolition of the Hindu head tax did not show political sagacity.
- o 26. In order to compensate his treasury for the deficit caused by the abolition of the Hindu head tax, Akbar instituted a religious pilgrimage tax.
- o 27. One of his greatest military victories was the battle of Panipat.
- x 28. Akbar's generals often criticized him for his generous treatment of the vanquished.
- o 29. Akbar died before Europeans came to India.
- x 30. Akbar built a new capital at Fatehpur Sikri.
- o 31. Akbar's greatest ambition was to unite all India under the religion of his fathers.
- x 32. Amongst Akbar's best friends were the two brothers, Abdul Fazl and Faizi.
- x 33. Faizi was a famous scholar of Hindu literature.
- x 34. Akbar's son Salim became his successor.
- o 35. Akbar lies buried at Fatehpur Sikri.
- o 36. As a successful general Akbar considered training in the army the best type of education for practically all men.
- o 37. The famous financier of Akbar's reign was named Birbal.
- x 38. Bengal and Orissa became part of Akbar's empire.
- x 39. Jahangir rebelled against his father Akbar.
- x 40. Akbar married a Rajput princess.

The advantages of a True-False test are several, most of which are shared by other New-Type tests. These will be discussed at length in Chapter IV, along with the limitations of New-Type tests. But any one who uses New-Type tests must know well their characteristics and how *not to use them*. The characteristics of the separate types, how to construct them and what pitfalls to avoid, will be discussed in this chapter briefly, deferring the broader question of advantages and limitations until later.

There are a number of characteristics of a True-False test which make themselves immediately apparent. One is that pupils can get a good many items correct by guess-work. But the guess-work has been provided for and a mere guesser will get approximately zero on this test for the simple reason that in a fairly long list of items the examiner assumes that out of every two times the examinee guesses, he accidentally guesses right once and wrong once. The examiner therefore infers that each wrong response means that two guesses have been made and deducts two marks to make up for the one got right by guess-work. For example, if a pupil got thirty items right and ten wrong, he would get ten marks deducted for wrong guesses and ten more for the accidentally right guesses, leaving a final score of twenty which denotes the number of items he knew positively [$40 - 10 - 10 = 20$; or, $40 - (10 \times 2) = 20$].

Let us assume that a pupil realizes that he knows practically nothing about the subject-matter, and decides to guess at all answers. He is quite likely to get twenty out of the forty items right. But twenty wrong and twenty more deducted for guessing, leaves a score of zero even though half the items were marked right. Usually in a test of this sort, the score recorded is the number right minus the number wrong, which is exactly the same thing as

deducting twice the number of mistakes from the number of items attempted. Questions not attempted are not counted. If only thirty out of the forty questions were attempted and all were right, the score would be 30.

It is quite fair to deduct two marks for every wrong answer for when the only alternative is to guess either 'right' or 'wrong', it can be assumed that in the long run about half will be guessed right and half wrong. The longer the test the more certain we may be that the above assumption is correct, for the law of averages applies to a large number of trials and is not at all reliable in a small number of trials.

When using a True-False test for an examination on which the result of a year's work is to be decided, there should be at least 100 items in the test to eliminate the unreliability arising from mere chance and guess-work. If a test is given largely for teaching purposes or for checking the work of the class as a whole, a shorter test can be used; but when promotion is at stake, the longer test is needed.

Another feature of this type of test is that it is time-saving and enjoyable to the pupils even though sharp thinking and exact knowledge may be demanded for a good score. It is also apparent that one can easily confuse the pupil who learns largely by rote and give opportunity to the pupil who can draw the proper conclusions and inferences from what he has learned. It is easy in such a test to steer clear of the stereotyped and bookish answers of the traditional test. One can test the application of learning in unexpected situations and see how generally available the new learning is to the pupil, rather than how well he can repeat the lessons word for word.

In constructing True-False tests, one should take care not to make the statements preponderantly negative or pre-

ponderantly positive. About a half-and-half proportion of positive and negative statements is desirable or the clever student may notice that the teacher consistently gives more of one kind than the other, and when in doubt as to the correct answer, will mark the way which gives him the better chance to guess right. If a pupil notices that his teacher usually gives twenty-five per cent negative items and seventy-five per cent positive, marking all doubtful ones positive increases his chances considerably. Symonds¹ says:—

Some of the pitfalls to be avoided in making a True-False statement are:—

- (a) Avoid long statements. Good statements can be constructed with from 10 to 20 words. Longer statements are sometimes necessary.
- (b) Avoid ambiguous statements.
- (c) Avoid trivial statements.
- (d) Avoid suggestive statements.
- (e) Put statements positively where possible.
- (f) See that one statement does not answer a preceding one.
- (g) Avoid dependent clauses (especially those which determine the truth or falsity of the statement).
- (h) Avoid compound sentences containing two separate ideas either of which may be false. If either half of a compound sentence is false, the whole item must be considered false.

Such papers do not take long to correct if one goes about it properly. A 'key' can be made by taking one of the test papers with all the responses recorded as expected in a perfect paper. This paper is held beside the pupils' papers and the signs on them compared.

Some objection has been raised to putting false information before children. It is claimed that the suggestion of

¹ Symonds, P. M., *Measurement in Secondary Education*, p. 26. The Macmillan Co., New York, 1930

misinformation may be strong enough to create a wrong association in the memory. However, life continually confronts us with the need of culling misinformation, half-truths, and biased propaganda from the truth and so a little training in the examination of statements for their truth or falsity will not come amiss. There is no question of practising deception for it is plainly stated that some statements are right and some are wrong, and the student is challenged to say which. The claim that misinformation is taught through suggestion has been tested but it has been found that the examinee gets more aid in correcting misinformation than harm from the suggestions of misinformation.

Of course pupils will make a few careless mistakes in tests of this sort, out of carelessness marking right what they know perfectly well to be wrong. Brief revision before handing in papers may eliminate a few careless mistakes. It has been found in questions of memory-recall that if we know the correct response with enough certainty not to require guess-work, the correct response to each statement comes with lightning rapidity. Hesitation usually means uncertainty, which necessitates guess-work, (excepting in novel questions demanding thinking). Our first answer is more likely to be right than the answer we give after toying around with the statement. When pupils are given the opportunity of checking their papers, the changes they make (exclusive of correcting mere careless mistakes) are more likely to be wrong than right. Thus the opportunity of checking is of doubtful benefit.

A philosopher has more trouble with designating a statement categorically right or wrong than a child has. He knows that there are two sides to many questions. The writer was first introduced to this type of test in his later student days when his studies were largely of a philosophi-

cal nature. He found it hard to give the unqualified responses demanded. The instructor put the matter before the class in this way: 'If your first reaction is that the statement is nearer right than wrong, then mark it right. If your first reaction is that the statement is nearer wrong than right, then mark it wrong.' When we enter into the game in this spirit the philosopher's objections are not so serious. Life demands similar decisions from us in highly important matters every day.

Pre-adolescent children are not troubled much by philosophical debates of this nature and children above that age can be made to understand, after a few such tests, that there is a value in using snap-judgement in this way and that the answer we give has only relative meaning.

The Yes-and-No test is a variation of the True-False with the items worded somewhat differently. Both True-False and Yes-and-No tests may be modified to offer a third alternative such as 'unknown' or 'uncertain'.

Examples

- | | | | |
|--|-----|---------|----|
| 1. Were Mongolians the first people to tame the horse? | Yes | Unknown | No |
| 2. Were carts used in India in the time of Asoka? | Yes | Unknown | No |
| 3. Was gunpowder first manufactured in China? | Yes | Unknown | No |

The correct alternative is to be underlined or ticked. 'Unknown' does not mean that the pupil is ignorant of what he is expected to know but that exact information on the point is unavailable.

2. THE MULTIPLE CHOICE TEST

The Multiple Choice test is a very popular one because it readily lends itself to a variety of demands, namely, memory for detailed information, practical reasoning ability, and ability to choose the correct solution from a

number of suggestions all of which may sound plausible enough to require keen discernment. Practically the same material covered in the specimen True-False test given on pp. 5-6 is covered in the following example

Directions

In the following statements four alternative endings to each sentence are given. Only one is right. Put a cross (x) before the one that is right. Mark only one alternative.

It is well to include one example actually worked in the directions for the sake of those pupils who have never worked such a test before; or the teacher should demonstrate on the blackboard.

1. Akbar was born in Sind instead of in the royal palace because
 - (a) His mother's home was there.
 - (b) Good maternity care was to be had there
 - (c) His parents were fleeing before Sher Shah.
 - (d) His parents were there on vacation.
2. Akbar permitted his regent, Bairam, to go on a pilgrimage to Mecca because
 - (a) Bairam had asked to go, for religious reasons.
 - (b) Akbar wanted him to go on a diplomatic mission.
 - (c) Bairam had been rebellious and therefore Akbar wanted him to absent himself from India for a while.
 - (d) Akbar could not go himself and therefore sent a substitute.
3. Akbar tried to found a new religion because
 - (a) He thought the old religions false.
 - (b) He wanted to unite the people of many religions who lived in his empire under one religion.
 - (c) He had a vision telling him to do so.
 - (d) He was anxious to become a famous religious hero.
4. Akbar is famous in a way that very few kings and emperors of India are because
 - (a) He was a great conqueror.
 - (b) He had the most splendour at his court.

- (c) He built the finest cities and buildings
- (d) He moulded the people of many religions, races, and allegiances into a well administered and contented empire.

5. Akbar abolished the Hindu capitation tax primarily because

- (a) The revenue derived from it amounted to very little.
- (b) It made the Hindus in his realm more contented when they were not discriminated against by such a tax.
- (c) He thought of more effective ways of raising revenue
- (d) He could not enforce its collection.

6. The famous finance minister of Akbar's reign was

- (a) Todar Mal.
- (b) Abdul Fazl.
- (c) The Raja of Jodhpur.
- (d) Bairam

7. When Akbar came to the throne the extent of his empire was

- (a) Larger than Babur's.
- (b) Larger than Asoka's.
- (c) Smaller than Babur's.
- (d) Larger than the present Indian Empire.

8. At Akbar's death the extent of the empire was

- (a) The same as Babur's.
- (b) Smaller than Babur's.
- (c) The same as the present Indian Empire.
- (d) The same as Asoka's.

9. Abdul Fazl was famous for

- (a) Conquering the Deccan
- (b) Heading a rebellion.
- (c) Writing a history of Akbar's reign.
- (d) Organizing land surveys.

10. Akbar reigned for

- (a) 25 years.
- (b) 36 years.
- (c) 10 years.
- (d) 49 years.

11. Akbar was succeeded by his son
 - (a) Jahangir.
 - (b) Shah Jahan.
 - (c) Humayun.
 - (d) Aurangzeb.
12. As a tribute to Akbar's organizing ability much remains even today of his
 - (a) Civil administration.
 - (b) Army organization.
 - (c) Religious organization.
 - (d) Organization of a department of education.
13. In the field of religion Akbar's greatest contribution was in his
 - (a) Founding of a new religion.
 - (b) Expansion of his father's religion.
 - (c) The building of temples or mosques.
 - (d) The religious toleration of his administration.
14. The revenue that Akbar was able to raise was
 - (a) greater than,
 - (b) smaller than,
 - (c) about the same as,that which the present administration raises from the same geographic area

It is quite evident that the chance element of guessing right in a Multiple Choice test is one in four when four alternatives are given and one in five if five alternatives are given. Guessing in such a test will produce only a 25 or 20 per cent result. Every three or four wrong answers, as the case may be, may be assumed to mean that one right response was obtained accidentally through guess-work.

In making this type of examination, the difficulty of the test is almost entirely dependent on what kind of alternatives the questioner prepares. A question like: 'In the year 1572 Akbar subdued (Australia, America, Iceland, Gujarat)' is obviously ridiculous and no test at all. The

COLUMN I	COLUMN II
<i>Causes</i>	<i>Effects</i>
7. Warm ocean streams	(7) Mild even climate
8. Dry climate	(11) Tidal action
9. Moist warm climate	(1) Freezing over of rivers and lakes
10. Volcanic activity	(8) Hot days and cold nights
11. Attractive force of the moon	() Typhoons
12. High humidity with a sudden fall of temperature	(3) The dry climate of Tibet
	() Rainbows
	(2) Cold climate

(The numbers in brackets under Column II have been filled in for purposes of illustration.)

The score is the number of correct matchings. Not more than fifteen to twenty items should be included in each column for if the number of items is more than this the pupil will have to re-read a long list too many times. One need only try to solve a Matching test oneself to see what is meant by this last statement. Of course one can have more than one Matching test in the same examination. No less than ten items should be included in at least one of the columns. The columns need not have the same number of items as some of the items in one column may remain unused, being put in to increase the number of possible choices for the examinee.

One may arrange for only one item from Column I to apply to each item in Column II or one can make a more complex test as in the following example in which each item in Column I has several possible correct associations in Column II, all of which should be indicated. In such a case it is best to inform the examinees that there are several possibilities and that the score depends on the number of those indicated.

COLUMN I	COLUMN II
1. Stork	(12, 13) plant
2. Bat	(3, 4, 9, 11, 14) invertebrate
3. Snail	(1, 2, 5, 6, 7, 8, 10) vertebrate
4. Butterfly	(5, 6) fish
5. Shark	(1, 10) bird
6. Herring	(2, 7, 8) mammal
7. Whale	(4, 9) insect
8. Hyena	
9. Beetle	
10. Myna	
11. Earthworm	
12. Fern	
13. Diatom	
14. Scorpion	

4. THE COMPLETION TEST

As its name implies, the Completion test is a test from which something has been left out, and this is to be supplied by the pupil. There is nothing particularly new about this type of test; the filling in of blanks in incomplete statements is an old pedagogical device. However, the modern testing movement has made much wider use of this way of putting what is in effect a question, and places much more emphasis on the fact that the statement must be so worded that the answer treated as 'correct' in the test is practically the only answer that can reasonably be defended.

A Completion test may also make use of pictures. Several intelligence tests for young children give a picture in which the eye or nose of man or animal are left out and ask the child: 'What is missing in this picture?' Pictures, diagrams and the like offer a large variety of possibilities for treatment in this way to test observation and understanding of relationships. We are here, however, dealing only with simple informal tests of the type

that the teacher can prepare for the classroom with no considerable expenditure of time and means, and so we shall confine ourselves to verbal Completion tests. The example is a test in language usage.

Directions

In each of the following sentences one word has been left out for each blank line. Think of the one word that should be written on the blank line to make the sentence correct and sensible. Write the word in the blank space. Read the whole sentence before you write the word. (Sometimes the first letter of the word you are to use is given.) Read the sentence again after you have written the word, to be sure that it is correct and sensible.

1. 'How much money have you, Rama?' Rama answered 'I ——— only three annas.'
2. 'Did you bring the parcel?' 'No, but I would have b—— the parcel if it had not ——— too heavy.'
3. 'W—— you surprised to see us?'
4. 'These books belong to Kashiram but th—— on the shelf are mine.'
5. The committee ——— composed of three members.
6. When we do our daily lessons faithfully we are more l—— to have satisfactory results.
7. 'Does the cycle belong to you and your brother together?' 'Yes, it belongs to my brother and ———.'
8. He is not as tall ——— you are.
9. When you finish your work in this room, please return to your ——— room.
10. Kamala was so sure she had won the prize that she asked the teacher: 'Who won the prize? Was it ———?'
11. 'W—— will you choose to be chairman of this meeting?'
12. Both Mohan and Jagadish are strong but Jagadish is the str——.
13. Jagadish is the strong—— man in the football team.
14. They let Phirtu play with them although he could not play nearly as well as ———.

15. The boys enjoyed diving. They dived —— a springboard —— the water.
16. The shopkeeper signed —— to the receipt.
17. Every scout is asked to bring —— own cooking vessels.
18. I would have gone if I c —— have made the proper arrangements.
19. 'Let us work together. You and —— will finish more quickly by helping each other.'

Symonds¹ gives the following suggestions for those preparing such an examination:

Care must be taken not to omit too many words as the test becomes difficult when the omissions are numerous. The best plan is to prepare a list of important statements covering the portion of the subject on which the pupils are to be tested and then from these statements to strike out a significant word or words. In every case the words omitted should be the only ones that will fit in the blank spaces. The pupil's score is the number of blank spaces filled in correctly. Oftentimes, even though great care is taken, a word other than the one stricken out will fit in the blank space. In this instance, the scorer will have to use his judgement as to whether it should receive credit. In giving a completion test each pupil must be provided with a mimeographed or printed copy of the examination.

The chance of 'guessing right' in this type of test is small if the statements are well worded. It is not an easy test to prepare and there are pitfalls for the inexperienced examiner.

5. CLASSIFICATION TEST

Directions

In each line cross out the word that does not belong there. Cross out just one word in each line, as for example:—

Cat, horse, cow, *mango*, dog

¹ Op. cit., p. 37.

1. Punjab, United Provinces, Rajputana, Bombay Presidency, *Kashmir*.
2. Narbada, *Bay of Bengal*, Ganges, Indus, Mahanadi.
3. Europe, Africa, Asia, *England*, North America.
4. Harbour, trade, exports, imports, *mountains*.
5. Tahsil, district, *island*, province, division.
6. *Lahore*, Bombay, Colombo, Madras, Karachi.
7. Nagpur, Lahore, Ludhiana, Amritsar, Rawalpindi.
8. Ceylon, Andaman, *Burma*, Laccadive, Nicobar.
9. Kashmir, Gwalior, Hyderabad, *Bengal*, Mysore.
10. Monsoon, rice growing, *desert*, fifty-inch rainfall, forests.
11. Peninsula, isthmus, plateau, island, *lake*.
12. Lake, ocean, *mountain*, river, sea.
13. Lake, river, *ocean*, pond, village tank.
14. Railway, *industry*, steamship line, highway, trade route.
15. Nilgiris, Himalayas, Vindhya, Western Ghats, *Indus*.
16. Nepal, *Ceylon*, Bhutan, Sikkim, Kashmir.
17. *English*, Gujarati, Hindi, Tamil, Telugu.
18. *Bangalore*, Pondicherry, Goa, Diu, Karikal.
19. Elephant, tiger, water buffalo, *kangaroo*, sambur.
20. Pipal, palm, tamarind, teak, *birch*.

(The words to be crossed out are printed in italics.)

At first sight this test seems to be much like the Multiple Choice test. Most of the rules which apply in making up a Multiple Choice test apply here also; so too the remarks about the element of chance in guessing correctly when the right answer is not positively known. But a careful examination of the mental process of doing this test shows this type to be quite different from the Multiple Choice. The Multiple Choice test puts several alternatives before us from which we must choose the right one. The Classification test on the other hand demands that we discover a common relationship between four out of five words which is not shared by the fifth. Our real job

is not only the choice of the word to be struck out, but a much more complicated process; namely, discovering the principle of classification which makes four of the words related and the fifth word an imposition into the series. This unrelated word can be identified as the unrelated word only after the principle of classification of the others has been discovered. We do not choose the right word, we strike out the nuisance which is quite a different mental process. This type of test is commonly called an XO test (Cross-out test).

This list of examples does not exhaust the variety of New-Type tests. By making use of pictures, diagrams, charts and so on, considerable variety can be introduced. But nearly all tests can be resolved into a few main types with numerous variations of each type. For use in the classroom by a teacher who has limited time for test construction and experiment and cannot afford to get tests printed or blocks for illustrations made, the five tests described will prove the most useful and are quite sufficient to do all that most teachers will want to do through their own informal tests.

The different types of test may be used in one and the same examination when once the pupils are familiar with the various types. If they are not familiar with the various types the different procedures required are liable to prove confusing. The following geography examination was composed 'in the rough' in less than two hours. Once the teacher 'gets the hang of it' it does not take long to make informal tests.

PART I

Directions

Some of the following statements are true and some are false. Put a cross (x) to the left of each *true* state-

ment and a zero (0) to the left of each statement that is partly or wholly *false*.

Do not mark the statements you do not know

1. The range of temperature in Colombo is greater than in Lahore.
2. The monsoon reaches Burma sooner than it reaches Calcutta.
3. Jacobabad receives the heaviest rainfall in India.
4. The climate of Asansol is better for cotton growing than that of Nagpur.
5. The Sundarbans country is south of Calcutta.
6. The Santal Parganas are in the vicinity of Poona.
7. One of the principal crops of Bengal is jute.
8. One of the principal crops of Bengal is cotton.
9. Nanda Devi is the highest mountain in British India.
10. Since Kashmir is a mountainous country, rice is not grown there.
11. The people living in the wheat-growing parts of India are on the whole taller than the people living in the rice-growing parts.
12. Lahore has heavier rainfall than Benares.
13. The aboriginal peoples of India are darker in colour than the people of the later invasions.
14. Bangalore is a sea port on the west coast of India.
15. Much rubber is produced in Ceylon.
16. When forests are cut down the flood conditions in the country become worse.
17. Since the best railway and boat connexion between England and India takes not less than 11 days of travel, one cannot expect a letter from London to arrive in Bombay in less than 11 days.
18. India carries on more trade with the United States of America than with England.
19. Marathi is the most widely spoken language of India.
20. Large quantities of lumber come from Kashmir to the Punjab every year. This is mostly teak wood.
21. In Ceylon the same postage stamps are not in use as in peninsular India.

22. The Buddhist religion flourishes best in the vicinity of the birthplace of the founder of this great religion.
23. Burma is no longer a political part of India
24. The manufactured products of India are of less importance economically than the agricultural products.
25. The ministers of each province are appointed by the Viceroy.

PART II

Directions

In each of the following sentences a word has been left out at the place of the dotted line. Fill in the correct word.

1. There is an important city in the Deccan called Hyderabad. There is another important city called Hyderabad in . . .
2. The rich grasslands of Gujarat have made it possible for this area to develop the. . . industry to such an extent that Gujarat is famous for the products of this industry.
3. Cawnpore is famous for its mills.
4. The Jumna and Ganges join near the city of.
5. . . . is often called 'the roof of the world'.
6. The mountain range in which Mount K2 is found is called the range.
7. Katmandu is the capital of
8. The number of provincial Congress Governments at present is
9. . . . is the most important river port in India.
10. The most important export of India to Japan is

PART III (MULTIPLE CHOICE)

Directions

In the following statements four alternative endings to each sentence are given. Put a cross (x) to the left of the one that is *right*. Mark only one alternative, the one you think is the best answer out of the four given.

1. Bombay is a more important port than Calcutta because

- (a) It is an older port.
 - (b) The harbour is larger.
 - (c) Bombay is closer to the centre of population.
 - (d) Bombay is closer to that part of the world with which India has closest commercial and political relations.
2. The export of indigo, which is used in the manufacture of certain dyes, is not nearly what it was 50 years ago because
- (a) Dyes like that made out of indigo can now be manufactured from chemicals.
 - (b) People do not like the colour indigo as much as formerly.
 - (c) Other countries grow better indigo.
 - (d) The entire indigo crop is needed for local use
3. The most thickly populated part of India is
- (a) Hyderabad.
 - (b) Bengal.
 - (c) Malabar.
 - (d) The Punjab.
4. Dehra Dun is not far from Ambala but has a much higher rainfall because
- (a) It is farther north.
 - (b) The monsoon gets there earlier.
 - (c) It is close to very high mountains which draw the rain.
 - (d) Rice is grown there.
5. There is no railway running into Tibet because
- (a) It takes much money to build a railway and the Tibetans are too poor to build one.
 - (b) There is no iron in Tibet from which to make railway lines.
 - (c) Tibet is such a small country that a railway is hardly necessary.
 - (d) High mountains surrounding this country add to the inaccessibility and isolation of Tibet.
6. Tea growers in India are having a hard time to find a market for all the tea they can produce because
- (a) The supply is greater than the present demand.
 - (b) Tea drinking has been found to be harmful.

- (c) The price of Indian tea is higher than that from other countries.
- (d) Indian tea does not compare favourably with that of other countries.

7. Bombay city does not become as cold as Nagpur because

- (a) It is further south.
- (b) It is nearer the sea which tempers the climate.
- (c) It is a bigger city.
- (d) It is farther from the Himalayas.

8. The southern slopes of the Himalayas are covered with trees. North of the high mountains very few trees are found. The reason for this is that

- (a) The soil is too rocky.
- (b) The altitude is too great.
- (c) The mountains prevent sufficient rain from falling north of the mountains.
- (d) The land is all cultivated, leaving no room for forest.

9. The main reason why Bengal grows much more rice than wheat is that

- (a) Bengalis like rice better than wheat
- (b) The price of wheat is higher than that of rice.
- (c) Rice is a healthier food than wheat.
- (d) The climate of Bengal is better suited to rice growing.

10. The great difference between the culture and institutions of the peoples of north India and south India is largely due to the fact of

- (a) A different climate.
- (b) A difference in geographical features of the country.
- (c) The difference in diet of the people.
- (d) The fact that north India is directly in the path of the various invaders of India from the north-west.

PART IV (MATCHING TEST)

Directions

In Column I is a list of cities. In Column II is a list of items pertaining to certain cities. Read over the list of items in Column II and see which applies best to

Karachi. Then write the number printed in front of 'Karachi' in the brackets in front of the item you choose as applying best to Karachi. Mark only one item, the one that fits best. Do the same with the name of each city.

COLUMN I	COLUMN II
1. Karachi	() Famous because of the Golden Temple of the Sikhs
2. Quetta	() Capital of one of the most progressive Indian states
3. Tatanagar	() Capital of Hyderabad
4. Amritsar	() Famous for oil production
5. Akola	() Famous for coal production
6. Lucknow	() Chief port for Sind
7. Mysore	() Irrawaddy river port
8. Mandalay	() City in Ceylon
9. Madura	() Famous because of one of the largest Hindu temples in existence
10. Goa	() Port for Malabar
	() Portuguese possession
	() Famous as a cotton mart
	() Famous for jute mills
	() Capital of the United Provinces
	() Recently visited by very severe earthquake

PART V

Directions

Opposite the name of each river, state in Column I the province, state, or foreign country in which the river rises. In Column II state the province in which the river joins the sea, or merges with a larger river of which it is a tributary.

COLUMN I	COLUMN II
Indus
Brahmaputra
Narbada

	COLUMN I	COLUMN II
Godavari
Jumna
Ganges
Mahanadi
Irrawaddy

A teacher can use the same test in succeeding years, or at least considerable parts of a test. If the pupils of a former class have no written copies in their possession, one can be certain that they will be able to give very little help to the present class that will be of value in the test. When there are half a dozen questions in an examination, one can sit down and record them from memory afterwards but not when there are from fifty to a hundred items. Thus a teacher will find the labour expended on a certain test this year of benefit in succeeding years also. In the course of three or four years he will have enough material at hand so that with a little reorganization and addition of fresh items he can quickly get up an acceptable test.

What is the passing mark on a New-Type test? In India we are so used to hearing of an arbitrary uniform percentage number as constituting the passing mark that it will be difficult for some teachers to adjust themselves to the fact that each test of the modern type has its own passing mark which no one can know until the test has been tried out to see how hard it is. If answer papers are arranged in order of excellence from best to poorest, one has the required information. Usually one considers that the twenty-five per cent of the pupils who are weakest are not doing passing work. But not always. There will always be the twenty-five per cent weakest pupils and sometimes these are nowhere near weak enough to justify our 'failing' all of them. Sometimes fifty per cent of the

pupils cannot 'make the grade', although in such a case probably the teacher or the test is at fault rather than the pupils. If four-fifths of the twenty-five per cent weakest pupils are just a little lower than the average mark for the whole class and the remaining one-fifth are far below, obviously the one-fifth (which is five per cent of the whole class) deserves to fail.

If sixty-five per cent of the pupils score well and thirty-five per cent are very much below the average mark the failures should probably be thirty-five per cent. There is no defensible rule of thumb for deciding passes and failures. Yet this question is decided by rule of thumb in most schools in India, and in most public examinations. But what an unsatisfactory state of affairs this has produced! Superficially it seems uniform and just to decide that fifty per cent of the total marks shall be considered passing, but when one year eighty per cent of the pupils of Class VIII pass in the English examination set by the Central Board and the next year forty per cent pass, we know it is not a matter of the pupils being so much stronger or weaker in one year than in another but of the examiners giving examinations that are not of uniform difficulty (which is inevitable, for nobody knows in advance how difficult a new examination will be) and making no adjustments to correct this. This is a most unscientific way of examining, which New-Type tests obviate by judging according to competitive standing. It is fairer to say that we will select such-and-such a percentage of the pupils standing best in the examination than to decide that every pupil who gets such-and-such a percentage of the marks in an examination (of which the degree of difficulty is pure guess-work on the part of the examiner) shall pass and that all those who get below a certain percentage of the marks will fail. The work of hundreds of pupils and

teachers is invariably a more uniform product from year to year than is a few hours work of an examining board which, because of the atmosphere of secrecy surrounding the examination, does not even get an opportunity to experiment with its questions to see whether they are easy or hard.

WHAT IS A STANDARDIZED TEST?

ALL the devices employed in informal New-Type tests, and many more, are utilized in Standardized tests. Standardized tests may be expected to be much more elaborate than the informal and certainly much more carefully and scientifically prepared. They are intended to be used in a large number of schools, with thousands of individuals. Very few teachers will ever prepare a Standardized test but practically all teachers in progressive schools may expect to give Standardized tests at some time or other. Through the use of Standardized tests a teacher can quickly compare both his own class as a whole and the individuals of his class with other classes and individuals. Information to enable him to make such comparisons is a necessary part of a Standardized test. A teacher must be well-informed in order to use such tests intelligently.

A fully-standardized test is standardized in three respects: (i) in form and construction; (ii) in the way it is administered and evaluated; and (iii) so that the score of any examinee can be quickly and easily compared with the scores of other examinees of the same age, school placement, or other classification. In other words, norms must be supplied.

Some tests are standardized only in one or two of these respects, and hence do not merit the name of fully-standardized tests. Even informal tests like those in Chapter I may be partly standardized. Ordinarily we expect a fully-standardized test to be comprehensive enough

to do a thorough job, its form and arrangement to be perfected through experiment, its material to be well graded, and results on the test to be available from not less than 2,000 examinees of various ages and accomplishments.

» The form and construction of a test may have a marked bearing on the results of the test. Supposing there are two editions of a reading test, one with large print, the other with very small print. The very small print may be disturbing enough to handicap seriously the group using it, especially in speed tests. Ambiguous directions or printers' errors also affect the result.

» In a printed test, much depends on the proper arrangement of the page. Suppose that in a Multiple Choice test each alternative choice is not clearly set off from the others, so that some pupils are confused, taking a comma which was meant merely to set off a phrase in one of the alternatives, to indicate a separate alternative. Or suppose again that in a Completion test some of the blank spaces to be filled in are so small and inconspicuous that some of the examinees fail to notice them. A well arranged and clear printing job that reduces the possibilities of confusion and misunderstanding on the part of the examinees is one of the first requirements of a Standardized test.

Any change in the form or content of a Standardized test is likely to change averages. In case of changes the test must always be tried out again to see what effect the changes produce in the scores. Seemingly small changes sometimes make a significant difference.

» The author of a test must be sure that the test is neither too hard nor too easy for the groups for which it is intended. Recently, in a reading test for middle school pupils, eleven items out of forty were found to be so easy that practically all middle school pupils could do them.

These eleven items had to be discarded and harder ones substituted. Ten items, on the other hand, were so hard that no pupils succeeded with them. These were therefore also replaced. All such experimental work needs to be done before the test can be considered adequately standardized. Usually authors arrange their questions in order of difficulty, judging difficulty by the ease with which examinees answer them. This is not done in all cases, since it is not always useful to have the material thus arranged.

There is also a difference in results according to what the examiner may say or do in the classroom. Once in administering the Goodenough 'Draw-a-Man' test, the writer said: 'I want you to draw for me the best possible picture you can of a man. Do your best and make it as good and as complete a picture as you can.' The teacher in the classroom, fearing that the children had not understood my poor pronunciation of Hindi, nobly came to my aid and re-explained. This was good except that he explained too much. He added: 'And don't forget to include hands, legs, mouth, nose, eyes.' That spoiled the taking of the test that day, for no suggestion must be given about what to put into the picture. The pictures of those who had had the advantage of hearing this list of suggestions could not honestly be compared with the pictures of those who had not. Many Standardized tests give the exact wording of everything the examiner is to say so as to be certain that all those taking the examination will have exactly the same amount of help, no more and no less.

Details as to the time to be allowed, and just how to work the different items, are printed on the test sheet or in a separate examiner's manual. These detailed directions the examiner must follow as closely as the man who uses dynamite for the first time might be expected to follow directions!

By no means all Standardized tests have printed forms for the examinee. The most famous Standardized test of all, the Binet-Simon intelligence test, has what amounts to a book of instructions for the examiner but no printed form for the examinee except a record sheet. To follow such elaborate instructions as are given in this test, a well-trained examiner is necessary. He must not only be well-grounded in psychology but must have practised giving the Binet-Simon test before he can do it well enough to get dependable results. Fortunately, the great majority of Standardized tests are far easier to administer, some of them being so easy that any ordinarily intelligent teacher who pays strict attention to the instructions given, and is scrupulously honest, can learn to give them. Tests for older people are fairly easy to give. There is a test called *The Otis Self-Administering Test of Mental Ability*¹ which, as its name implies, can be given by the examinee himself, so complete and clear are the instructions. It is assumed in this test that the examinee will administer the test honestly as he can cheat no one but himself.

The scoring or evaluating is also standardized. It is explicitly stated just which answers are acceptable and which are not. There can be no dispute on whether the answer is right or wrong for the author of the test has settled the question with finality. One may question the author's judgement privately but in marking the papers one must follow the author's directions or else the results of the papers are not comparable with the results of papers that some one else corrects.

In such tests as the True-False, only two answers are possible, true and false. This makes the test strictly ob-

¹ By A. S. Otis. Published by the New World Book Company, Yonkers-on-Hudson, New York.

jective, i.e., not subject to the examiner's personal idiosyncrasies, but according to the same objective standard of judgement for all examiners using that test. From this feature of the New-Type tests, they are often called 'objective tests' which name is used by many psychologists instead of 'new-type' or 'modern' which must obviously soon be out of date.

In the Multiple Choice test, the scoring is also strictly objective. In the Completion test it is difficult to frame all statements so that a single insertion is correct but with care Completion tests also can be quite objective.

Most Standardized tests are so strictly objective in the sense that the judgement and opinion of the evaluator does not come into consideration, that the scoring is simply a matter of seeing that certain words or figures are underlined, crossed out, or checked in the proper places. Sometimes a hundred pencil marks, such as checks, crosses, etc., are all that the examinee has written in an examination that covers as much ground as a hundred-page note-book would cover in the essay-type of examination. Recently a machine has been invented for evaluating papers. An electric current passing through the test paper reveals, by magnetic action on those spots to which pencil graphite has been applied, whether or not the pencil mark is in the right place or not. Such machines are said to be able to mark thousands of papers in a day and to do it far more accurately than any human being can be expected to.

That is the height of objectiveness in evaluating papers. Certainly no examinee can claim that a poor breakfast or a bad night's sleep on the part of the evaluator was responsible for his low marks. In the case of the human evaluator of the old-type examination one could never be certain that indigestion did not influence the score!

Recently a paper of arithmetic problems of the old-type was given to *ten professional* examiners to evaluate. Of a possible total of fifteen marks, one examiner gave 15, three gave 12, two gave 8, two gave 7, and two gave 2 marks.¹ This is by no means an isolated case and stands in striking contrast to the objective marking of a Standardized test.

The third and last respect in which a Standardized test is standardized is that it gives the standard of performance of the average boy or girl of specific ages, school grade, etc. For instance, in a certain arithmetic test it was found that the following were the average scores of the pupils in the various classes. (From 500 to 2000 pupils of each class had been examined.)

CLASS	SCORE
IV	20
V	24
VI	28
VII	33
VIII	38
IX	40

Classified according to age the pupils got the following average scores:

AGE	SCORE
10	28
11	27
12	29
13	32
14	35
15	40

These statistics are called the 'norms' of the test. Each test has its own set of norms. The norm means the performance of the normal or average person. The use of

¹ Hartog, Sir Philip, and Rhodes, E. C., *An Examination of Examinations*, p. 24. Macmillan, 1936

this information is quite apparent. In order to find out if a pupil is above normal or below normal we need only compare his score to the norms. If Mohan, who is fifteen years old and in class V, gets a score of 24, we see that he is normal for his class but below normal for his age. If headmaster Tiwari finds that the average score for his sixth class is 38, he may be justly proud for this is the norm for the average eighth class. If in addition, he finds that the average age of his class is 12, he has a further reason for feeling proud because his pupils are equally superior compared to others of their age. No Standardized test is fully standardized until it supplies information of this sort, and an author of a test is in honour bound to point out the incompleteness of a Standardized test that has as yet no norms.

The sets of norms most used are the age norms and the school-grade norms. Sometimes a different set of norms is required. Suppose in some ability there should happen to be a considerable difference between the sexes. The sex norms would then also be useful. In the case of some psychological tests these are given although in school subjects there has not been any consistent difference between the sexes large enough to be taken into general consideration. In athletics, height and weight norms may prove useful. There is talk of having rural and urban norms, for in certain things the city children do have an indisputable advantage. In India, where the city monopolizes such a large percentage of the educated groups and the rural areas are backward, a different standard will probably be considered only natural when we realize the handicaps of rural children. This will call for different norms in Standardized tests for the two groups. In some schools in America a higher standard is expected from the intelligent pupil than from the average or dull. In

effect this means separate norms according to intelligence quotient.

In intelligence testing the age norms are the more important. In the testing of school subjects the class norms are more important.

There are two qualities a good standard test must have. It must be *reliable* and it must be *valid*. These two terms, 'reliability' and 'validity', have acquired a technical meaning which is something more than the usual meanings of these words. Both these qualities are measurable.

These two technical terms are not easy to understand but since they are used much in the discussion of tests, an attempt to illustrate their essential features should be included here. Some readers may prefer to skip the rest of this chapter for the time being and refer to it after the chapter on statistics.

The writer recently had occasion to see a very unreliable primary-certificate test given. In a certain class of twenty-four boys only eight had passed in arithmetic. The inspecting officer consented to give another test, which was of the same difficulty as the first. This time seven boys passed; but out of these seven, only one was from among those that had passed the first test. Thereupon a dispute arose between the parents as to which test should be considered decisive. As the results showed, there was approximately as much chance for the boys of the class as a whole to pass the one as to pass the other. But certain individuals found it much easier to pass one than the other. The two tests seemed to be of like difficulty but the results on either or both of the tests were certainly not reliable measures, for the results of two-thirds of the pupils were entirely different on one test than on the other. Were both of these examinations a reliable measure, the one who stood first in the one would also

have stood first in the other, the one who stood second in the first would also have stood second in the other, and so on.

The reliability rating (or *consistency* rating as Sir Philip Hartog calls it) of a test is usually made in this way. Assume that a test was given on a certain date. One month later, when the pupils have forgotten all about what was in the test, the examiner comes along, and to the surprise of the pupils, gives them the same test over again. The examiner arranges the papers or scores of both tests in order of merit and sees to what extent those who stood at the top of the class still stand there and those who stood at the bottom of the class still have that negative distinction. If the order is approximately the same (one cannot expect it to be exactly the same) we may safely assume that the test is *reliable* or *consistent*, for it does today what it did a month ago. In the meantime the pupils may have learned a number of things and bettered their scores but that will not alter their relative standing. If on the other hand, some who were high are now low and many who were at the bottom of the list are now at the top, we know that chance plays a large part in the score and that the test is *unreliable* or *inconsistent*. Reliability means that the score of a pupil on a certain test is not a chance good score but a consistent good score, or that a pupil's work is not chance poor work but consistent poor work.

Another way to find out if a test is reliable is to separate the scores on the odd-numbered items and treat them as a complete test to be compared with the scores on the even-numbered items, given as a second test. Thus you try one half of the test against the other half. If the pupils have the same relative standing on the one half as on the other, we know that they do not get the good

or poor marks by mere chance. Or, if there is already a test that is known to be reliable one may compare results on it with results on the new test of the same kind. This will also give a measure of its reliability.

The amount of reliability can be worked out through treatment with a proper mathematical formula of the data given in a 'test and re-test' or 'split-halves' test such as is described in the two preceding paragraphs. The method of working out the mathematical expression is rather complicated and requires some understanding of statistical mathematics. Since the mathematical method is the same as that involved in working out the correlation of two tests, that part of Chapter X which treats of the mathematics of 'correlation' applies to the mathematics of a reliability test also.

Perfect reliability is expressed by the figure one (1.00). Absence of any reliability whatever is expressed by the figure zero (0.00). In between these two quantities we have all possible gradations between perfect reliability and zero reliability. In practice we do not expect to find perfect reliability. But some tests do boast of a reliability of 0.95 although such consistency is unusual. A good Standardized test should have a reliability of not less than 0.70, and 0.85 is considered a fair reliability. Usually the longer the test the more reliable it is, but this is not invariably the case. One cannot expect the highest reliability ratings where the examinees are not used to New-Type tests since consistency in performance depends on examinees knowing exactly what is expected of them. An examinee who took the same test twice without understanding how to work it in the first trial whereas he did understand in the second trial will perform very inconsistently on the two tests. A reliability of from 0.65 to 0.85 is probably not too low in India, especially if the

examinees have had little schooling and training in taking New-Type tests

Recently six primary-certificate tests in arithmetic were found to have a reliability, when one was considered an equivalent test for the other, of from 0.29 to 0.45. This reliability is so low that it reveals the score on these tests to be quite dependent on chance. Tests with no more than ten questions are bound to be unreliable because so much depends on the selection of the few items included. Pupils usually call a question to which they happen to know the answer 'easy' and one to which they do not know the answer 'hard'. The more questions there are, the more one can expect that every pupil will show what he really can do and the more confident one can be that he was not merely lucky in getting the questions he happened to know or unfortunate in getting just the few that he was least prepared to do.

When is a test valid? A test may be highly *reliable* and yet not be at all *valid*. I once saw the Hindi translation of a well-known English Standardized test in arithmetic problems. The pupils did very poorly in the test and the examiner decided they were poor in arithmetic. On looking at the translation however, it was soon plain that very high-flown Hindi, such as the children seldom heard or read, was used. It was found that in every problem there were from four to six words which the majority of pupils did not understand. So this test was a rather stiff vocabulary and reading test as well as an arithmetic test and could not be considered a valid measure of arithmetical ability. It was found that by rewording the problems the results were much improved. A test of low validity is a test that claims to test one thing but really tests another, a situation which makes it about as useless as the judgement of a doctor who mistakes malaria for diabetes.

There are many tests which claim to test intelligence but to discover whether they actually do is most difficult. Intelligence is not easy to define, let alone measure with exactitude, although some efforts to do so have undoubtedly met with notable success. If an alleged Intelligence test gives greatest credit to those who prove themselves in later life to be rank blockheads, it is obviously not an Intelligence test. If an Arithmetic test gives highest honours to those who later prove themselves to have little number sense and poor computational ability, or if a Reading test does not actually differentiate good readers from poor readers, they are not what they claim to be, and hence not valid as Arithmetic or Reading tests.

The real validity of a test is its proved usefulness. The Army Intelligence Tests used in the Great European War proved their validity by doing fairly well what they were intended to do, namely pick out the best men for responsible positions requiring intelligence. Mechanical-aptitude tests are valid if they can predict who will become a good mechanic, assuming that the examinees do their best to become good mechanics.

To determine whether a test is valid or not takes experience and time. Certain things can be done even during the experimental stage that give at least some indication of future success or failure of the test to do what it should do. If a new Intelligence test shows practically the same results as several other well-tried tests, we may safely assume that it does somewhat the same thing as the well-established tests do. One of the first things to do in finding out if a test is valid is to compare its results with those of tests that have given evidence of being valid. The extent to which it does this can be determined mathematically in the same way that the reliability of a test is expressed. If the results of a test also correspond fairly

well with the opinion of a reliable teacher as to the relative ability of those tested, this is another point in its favour although teachers make many mistakes in estimating the relative native ability of the pupils under them. School progress also is a criterion and any test, in subjects that are taught over a period of years is expected to show norms indicating steady increase in the ability of the pupils as they reach the higher classes; or the validity of the test may be rightly questioned. The norms of an Intelligence test must show development from class to class because one does expect 1000 pupils of Class VI to display more intelligence than 1000 pupils of Class V in almost anything they attempt to do. This last statement does not apply in the case of tests in school subjects where memory of recently drilled subject-matter, is greatly stressed, because in work of this type forgetting soon sets in and pupils will not make as good a score as when the learning is recent.

A Standardized test is a finished product, a work of art. It must conform to exacting standards of form and arrangement and approximate to a yardstick in the objective way in which it is applied and the results recorded. It must also prove to be a reliable measure which measures just what it claims to measure.

III

SCHOOL-SUBJECT TESTS (STANDARDIZED)

Most Standardized tests come under two main classifications: School-subject tests or Psychological tests. School-subject tests grow out of the school routine and examine in those things purposely taught in school, namely the knowledges, skills, and arts. Psychological tests on the other hand aim at the measurement of attitudes, emotions, natural ability, temperament, etc. Foremost and best-known among the Psychological tests are the Intelligence tests which must not be confused with School-subject tests. School-subject tests will be discussed at length before taking up the question of Psychological tests. Intelligence tests attracted public attention before Standardized School-subject tests, and were largely instrumental in stimulating the development of the latter. But as this book is intended for the teacher rather than for the psychologist, we shall first consider the School-subject test because it operates in a field in which the schoolmaster is more at home.

We often hear the terms *Scholastic* test and *Achievement* test used interchangeably with subject tests. The name Achievement test is usually reserved for a *series* of School-subject tests combined into a single battery of tests to measure all-round achievement, but the distinction is not very important nor is it consistently made, so the reader may regard the above three names as being synonymous. The purpose of these tests is quite different

from that of the Intelligence test although in form and partly in content they may resemble some Intelligence tests very closely. These tests deal with what the teacher tries to teach while Intelligence tests try to avoid teaching effects as much as possible. This is a very important distinction.

The school has been from ancient times, famous for the teaching of 'the three Rs': reading, writing and arithmetic. It is fitting that we start our description of School-subject tests with an examination of these three specialities of the school. Hardly any other three subjects could better illustrate the problems of scientific testing.

1. READING TESTS

In preparing Standardized Reading tests we must first ask what we are aiming to teach the pupil to do when we teach him 'to read'. We then set about trying to measure the extent to which pupils are learning to do this. First we must know what the aims of teaching reading are or we cannot arrange to test the progress in attaining these aims.

The traditional reading examination, not only in India but nearly everywhere, consisted of oral reading. By this method we find out whether the pupil, with only such aids as looking at the printed page, can repeat vocally the contents of the page exactly as printed.

This method of testing reading has some obvious weaknesses. First of all it tests the kind of reading that, outside the school, we need to do least of all; namely, reading aloud. A good reader does most of his reading silently, which is quite a different thing from reading orally. In oral reading neither the speed nor the accuracy of comprehension can be attained that can be attained in silent reading. To be able to do the one does not mean that we

can do the other satisfactorily. They both need to be taught and examined separately.

The traditional Reading test in India usually tests the reading of familiar material. This is no real test of reading for one cannot tell what part of the performance is real reading and what part mere rote memory. One cannot tell whether the reader reads understandingly or not from his reading an old lesson, for if he understood the lesson last week when it was taught, he understood it before starting his present reading performance. The conditions under which most oral reading is tested are artificial, for reading is of value mostly in the reading of unfamiliar material. Mere rehearsal of material previously read is of some value but not of much. Oral reading lends itself well to examining pronunciation, and accuracy of reading isolated words, but little else. For examining the understanding of what has been read, silent reading tests are much superior to oral. Understanding what is read is by far the most important thing about reading and can be entirely lacking in a person who reads aloud fluently and with fair pronunciation and expression. The testing of oral reading confines itself almost entirely to accuracy in vision and what might be termed the elocutionary effects. Since this can only be done by listening to each reader individually most Standardized tests omit oral reading, leaving this to the individual teacher to arrange in his own way. To those who wish to examine a Standardized test in oral reading, Ballard's *One Minute Reading Scale*¹ and Gray's *Oral Reading Tests and Check Tests*² are recommended.

¹ Ballard, P. B., *The New Examiner*, p. 145; and also in *Mental Tests*. University of London Press (see bibliography).

² Gray, Wm., *Standardized Oral Reading Check Tests and Standardized Oral Reading Paragraphs Test*. Public School Publishing Co., Bloomington, Ill., U.S.A.

Ballard's *One Minute Reading Scale* is an eminently practical little test which can easily be emulated in any Indian language. The test consists of a printed list of short words. The words are isolated and not related. The pupil is told to read these words aloud as fast as he can. At the end of one minute the examiner stops him and records the number of words read *correctly*. If the pupil hesitates more than five seconds over any word, he is told to go on. Dr Ballard supplies age norms so that the pupil's 'reading age' may be quickly ascertained. He supplies separate norms for boys and girls as he has found that girls read slightly faster than boys. This test tests only a single phase of reading, word recognition. It does not examine understanding or even pronunciation. There is a high correlation between scores on this scale and other tests which measure much broader aspects of the reading skill but the scale itself is a measurement of only one very narrow skill.

A similar list can easily be prepared in any language and used for comparison of pupils. Of course Dr Ballard's norms cannot be used for any list except his own. Such a reading test may seem at first to be nothing much more than the old-style reading examination; but note that it is very radically different for it tests material that is not at all familiar (a list of isolated words cannot be easily remembered like sentences where one word suggests the next), the scoring is quite objective, and the standards have been arrived at by careful experiment. This is an easy test which any intelligent teacher can work out for his own school, or preferably, a number of schools. The teacher is advised to use class norms rather than age norms as ages vary so enormously in each class of most Indian schools. A part of the test is included for illustration:

is	me	on	at	by	so	us	an	it	or	be
to	as	he	of	in	go	up	am	if	no	we
my	ox	do	the	end	for	but	him			
are	can	she	dog	let	you	not	was			
out	try	see	mix	cat	now	boy	saw			
bit	met	top	run	man	pet	lot	get			
did	van	bad	red	cup	bee	lit	pin			
had	ran	pen	nut	big	old	yet	rob			
gun	leg	fun	lip	new	fog	has	sit			
sly	wig	mud	box	ink	sat	end	cut			

The testing of silent reading has done much to revolutionize the teaching of reading. The emphasis is now on understanding and to some extent on the speed of reading, for it has been found that both accuracy and enjoyment of reading are greatly dependent on reading with ease and speed. It has been found that training can bring about improvements that were hardly considered possible thirty years ago. Probably half the students in high school and college could be taught to double their reading speed after a few weeks of proper training. It was formerly assumed that somehow or other those who got through the first stages of learning to read would advance to proper efficiency if they only kept on reading. The result of this unwarranted assumption has been, and is, inadequate readers in high schools, colleges, and in adult life—nearly all of whom could have been greatly improved with a little proper attention. The methods of examining reading reflect the kind of teaching of reading that experts wish to encourage.

The ability to read one kind of material does not give one ability to read all kinds of material. Recently I picked up a treatise on geology which looked interesting and undoubtedly is interesting to those who know how to read geology. But the terminology was not familiar to me and I soon lost interest because I did not sufficiently under-

stand. By reading geological subject-matter I will, however, in time acquire the ability to understand works of this nature. Some people read stories with interest but would be lost if given ordinary political news to read, or easy popular science such as ten-year-old children commonly read. What is easy for some people to read is difficult for others because they are not sufficiently familiar with the subject-matter and the particular vocabulary. In a reading test for middle and high school pupils, highly technical material should not be included but a fair variety of types of literature should certainly be represented, for the cultural value of reading is dependent on our ability to read as wide a variety of worth-while literature as possible.

The aims in the teaching of reading as stated in the *Twenty-fourth Yearbook of the National Society for the Study of Education*¹ have become classic and are accepted far and wide. These aims are briefly:—

1. Rich and varied experiences of reading.
2. Strong motives for and permanent interest in reading.
3. Desirable attitudes and economical and effective habits and skills of reading.

A test that aims at finding out how teacher and pupils are progressing towards these aims will of necessity embody features radically different from the old type of oral examination.²

¹ Published by the Public School Publishing Co., 1925.

² For a much more thorough discussion of the problem of teaching reading and measuring reading ability the reader is referred to the author's book *Suggestions for the Improvement of Reading in Indian Secondary Schools* (Oxford University Press), now in the press. In this book numerous test models, both formal and informal, are given in detail. Vocabulary tests are also included.

As an example to show the comprehensiveness of a single standardized reading test, an outline of the *Sangren-Woody Reading Test*¹ will be given. This was prepared for classes IV-VIII.

PART I. *Fact Material*

Forty multiple-choice items of the following type are given.

- A letter comes in the (water, wind, store, mail).
- A rock is a big (stone, animal, train, flower).
- (The proper word is to be underlined).

PART II. *Rate of Reading*

The pupil is given an interesting article on how a pencil is made. The pupil is given one minute to read as fast as he can, reading carefully however, as he will be tested later to see what he has understood. At the end of a minute the pupil marks on the page the place which he reached, and the examiner later records the number of words read per minute.

PART III. *Fact Material*

Six paragraphs with questions like the following are given. Note that the emphasis here is on detail. This is reading for detailed information.

The 'lead' in your pencil is not made of lead. Long ago people had lead in their pencils; that is probably why the pencils we use are called lead pencils. Another mineral called 'graphite' is now used. This mineral is taken from mines in the same way as coal or iron ore.

1. What did people use in their pencils long ago?
2. What mineral is used in pencils now?
3. From what is the mineral taken?

¹ By P. V. Sangren and C. Woody. Published by the World Book Company

PART IV. *Total Meaning*

The following paragraph does not aim at supplying detailed information to be remembered but rather at creating an effect. A word picture is painted which should be seen as a whole rather than in detail. This is a different type of reading from Part III. Twelve such paragraphs are given.

The waves rolled mountain high. They cut and slashed the sides of the ship. The wind howled and hissed like the voice of the angry sun-god himself. The heavy hail beat upon the deck like stones. Lightning struck the doomed ship and tore it into a thousand pieces.

What do the sentences describe? Summer, storm, mountain, rain.

PART V *Central Thought*

In paragraphs under this heading the author is testing the ability to get one central thought.

Cod fishing is a dangerous business. Thick clouds of fog often settle over the sea more quietly and quickly than showers of rain come down on us. In the thick fog, where one cannot see fifty feet, the men in the dories sometimes lose their way. Sometimes great storms come up and upset their little boats.

- (a) Cod fishing is a very dangerous business.
- (b) Heavy fogs often settle over the sea
- (c) Men cannot see fifty feet in the heavy fog.
- (d) Storms often upset the little boats.

(The pupil is to check the statement which is the central thought of the above paragraph.)

PART VI. *Following Directions*

Opposite are two squares of different sizes. The larger square is a playground for children, and the smaller one is a garden into which children must not go. There should be a fence between the playground and the

garden. Make this fence by drawing a line to separate the squares.



PART VII *Organisation*

Under this heading, quite lengthy selections are given and the pupil is required to organize the contents into an outline.

This Sangren-Woody test is a very elaborate one and is more than is required for ordinary use. There are many reading tests which are very much more simple and can be given in half an hour.

One of the simplest and most interesting ones is Ballard's which is given in his book *Mental Tests* (see Bibliography). This test consists of a single story, approximately 500 words in length. The story is an exquisite one about a small bird, a fish and a crow. The beginning of the story is very easy reading matter but towards the end of the story the reading becomes difficult and demands increasing maturity. The pupils are given three minutes to read the story, after which the printed story is taken from them and a completion test on the entire story is given them. There are 67 blanks to be filled in. This test gives a fairly accurate idea of the ability of the examinee to understand what is read and his ability to remember enough detail to make the reading profitable. A criticism commonly levelled at the test is that it gives too much credit for a good memory. But a reader who forgets as soon as he reads is hardly a good reader and ability to remember depends very largely on the extent to which the subject matter is understood. With the kind permission of Dr Ballard and the University of London Press, the

author of this book has adapted this test to Hindi and found it very usable in high school and normal school.¹

Any teacher can make up an informal Silent Reading test giving several paragraphs with appropriate questions on the meaning of each paragraph.

The writer has standardized five Reading tests in the Hindi language,² one of which is also to be had in Urdu. Part I of a test for middle schools contains nineteen paragraphs on which questions are put to test understanding. The second part is a vocabulary test (multiple-choice type). Vocabulary tests are often included with reading because a good vocabulary is essential to understanding. Part III is a simple 'rate of reading' exercise.

Two of the reading tests are for use in primary schools. Since the primary school concerns itself with the beginning stages of reading, the technique of testing these stages is quite different from that outlined above. A detailed description will indicate these differences. Norms are supplied for all the tests.

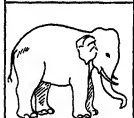
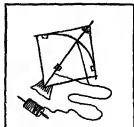
PART I. *Word Pictures*

Part I deals with the recognition of words. Several pictures are given, also several printed words. The pupil is told to draw a line from each word to the picture which illustrates the word. Thus pupils who can read the word 'elephant' have no difficulty in drawing a line from the word to the picture of the elephant, which is positive proof that they can read the word, excepting such cases

¹ See also the 'Python and Crow Reading Test' in the appendix of *Suggestions for the Improvement of Reading in Indian Secondary Schools* (In the press)

² These and other tests by the author subsequently mentioned are obtainable directly from the author, Bismampur, C.P., via Bhatapara

where the pupil has 'guessed' right. No pupil can guess right often enough to get a satisfactory score. Mathematics are against him. It is remarkable how many lines are drawn without a single one having been guessed right.



ELEPHANT

HORSE

HOUSE



The answers are quite obvious. The pupil is given one point for each word or phrase from which a line is drawn to the picture of it. The line may be drawn from any part of the word or phrase to any part of the picture, but

there must be no doubt as to which word or phrase and which picture is meant. In case of doubt, no score is given.

PART II. *Speed Test*

The pupil is given exactly three minutes to answer 'Yes' or 'No' to twenty questions which are so easy and self-evident that he should know the correct answers without much thinking. But in order to give the right answer he must be able to read the question correctly. If he does not read it correctly, his answer will be guess-work. He gives his answer by simply underlining the word 'Yes' or 'No', whichever is correct.

Here are samples of the type of questions.

Can a dog run?	Yes	No
Can you fly?	Yes	No
Have you four legs?	Yes	No
Is two more than four?	Yes	No
Can ducks swim?	Yes	No

(The correct answer is underlined by the pupil.)

By seeing how many such questions the reader can answer in three minutes we get some idea of his speed of reading understandingly.

The score is the number of right responses minus the number of wrong.

PART III. *Following Directions*

Part III consists of following printed directions.

Draw a tail on the dog



Make a cross mark in the circle.



Draw a circle around the bird.



Draw two lines under the monkey and mark a cross over the goat.

PART IV. *Paragraph Reading*

Part IV tests the ability to get the meaning of a paragraph as a whole, as in the following samples.

1. I am large.
I am still growing.

My leaves are green.

My branches are long.

Tell me, who am I?

The answer is one of the four words underneath.
Underline the right word.

girl house tree mango

2 Mohan went to the bazaar to buy a coconut. The shopkeeper asked six pice for a coconut. Mohan offered him an anna but the shopkeeper refused to sell. Mohan then went to another shopkeeper who offered a coconut for one anna but Mohan refused to pay more than three pice. A third shopkeeper offered the desired coconut for three pice but now Mohan insisted on getting it for two pice. So Mohan decided to climb into a nearby coconut tree and pluck himself a nut. Unfortunately he fell out of the tree and broke his neck.

- (i) How much money did Mohan give for a coconut?
- (ii) How many coconuts did he buy?
- (iii) How did Mohan get hurt?

As far as is known to the writer, there are as yet no other standardized reading tests in any Indian language although informal reading tests in the new-type style are being used regularly in quite a number of progressive schools.

Both Cyril Burt and P. B. Ballard have composed good reading tests in English which are obtainable from the University of London Press.

2. WRITING SCALES

Anyone can see that the testing of writing is quite different from the testing of reading. What we look for in writing is legibility. The test of this is in the reading. Some people may like fancy writing but the busy world has settled down to looking upon writing as a means of recording something that is to be read. So beauty and ornamentation in writing are considered not only less im-

विश्रामपुर लेखन माप ।

हिन्दुरथान मे तीन मौसिम होते हे । गर्मी के मौसिम मे खेत कट जाते हे, और अनाज गाहा जाता हे इस मे शादियां होती हे । खेतों मे खाद भी डाला जाता हे, और परो को मरम्मत होती हे

हिन्दुस्थान मे तीन मौसिम होते हे । गर्मी के दिनों मे खेत कट जाते हे और अनाज गाहा जाता हे । उस मौसिम मे, शादिया होती हैं खेतों मे खाद भी डाला जाता हे ।

portant than a neat, legible, everyday business hand, but out of place in correspondence. Fancy writing belongs to the field of art and not practical writing.

We all judge handwriting automatically as soon as we see it. We immediately classify it as *easy to read*, *fair*, *hard to read* or *impossible*. A writing scale does not do a great deal more than this, but does it with a little more detail and according to a uniform and consistent standard.

Most teachers can judge handwriting with some degree of fairness. Nevertheless they will be surprised if they evaluate one hundred specimens of handwriting and afterwards put together various specimens given the same grade. They will be surprised at how uneven in quality those given the same grade are. They will be equally surprised if they pick out ten specimens of the same quality and then see what varying marks they assigned to them. One cannot grade many specimens without letting one's standard of grading get more strict or more lax from time to time. A writing scale fixes the standard so that it varies relatively little.

The illustration facing this page is nearly self-explanatory. Even those who do not read Hindi can see that the first specimen is better than the second, the second better than the third, and so on down the scale. The examiner need only compare the writing specimen he wishes to evaluate with the specimen on the scale it most nearly resembles and assign the number given in the scale specimen. By having definite models of different grades of excellence before him as he works, the examiner's standard is relatively consistent. Even children can be taught to use this scale with enough accuracy for practical purposes and can thus keep record of their own progress.

A scale of this sort also serves the purpose of enabling

the teacher to keep a record of the pupil's progress of retrogression. Within the period of the school year most children's handwriting changes quite noticeably and not always for the better. By recording the pupil's scores at the beginning of the year the teacher can quickly inform himself as to the change taking place.

This scale is modelled after the famous *Ayers Handwriting Scale*¹ in America.

The construction of a scale is not as simple as it looks. One must be certain that the specimen labelled 'best' is really best, that the one labelled second best is distinctly better than the one immediately below it and also distinctly inferior to the one above it. The judgement of one person does not suffice to decide this. Two hundred teachers were consulted in the making of this scale and the scale had to be revised a number of times until there was a fair unanimity of judgement. Anyone interested in constructing a writing scale for any Indian language may find an article in *Christian Education*² for April 1937 of use. There is no point in constructing an English scale for use in India as almost any depot for standardized tests will be able to supply handwriting scales in English.³

Some constructors of scales give as high as twenty scaled models. But the difference in degree of excellence of the nearest specimens is so slight that it takes a writing expert to see the difference. Nine or ten specimens are quite enough in a scale for the average teacher and even more so when the pupil is expected to try his skill in judging.

A writing scale also comes into the category of a Standardized test, for the construction and method of its use

¹ Public School Publishing Company.

² The Lucknow Publishing House, Lucknow.

³ An English writing scale with four different styles of writing is given in Ballard's *New Examiner*, pp. 250-265.

is standardized, and norms are supplied for each class separately. Although it does depend on subjective judgement, it overcomes, in part at least, the idiosyncrasies of individual judgement by taking the consensus of opinion of a large number of people.

Some scales supply speed standards also, or suggest that certain standards should be attained at specified speeds. It is highly desirable to establish norms for speed in handwriting because in real life much writing has to be done 'against time'. A beautiful handwriting that is beautiful only when the writer writes very slowly is not sufficiently practical. Most schools ignore the need for reasonable speed in the penmanship hour thus causing a large gap between the standard in the penmanship class* and out of it. Speed should be made an objective in scales in order to restore a proper emphasis.

The Hindi scale illustrated does not supply speed standards as it is for use in the primary school where speed should not be stressed.

Prof N. C. Chatterjee of Patna Training College has composed a more elaborate writing scale in Hindi for use in middle and high schools. It contains twice as many gradations as the one illustrated for use in primary schools and also gives speed standards. Mr W. M. Ryburn of Kharar is working on an Urdu writing scale.

3. ARITHMETIC TESTS

A standardized arithmetic test is not unlike a traditional arithmetic examination except that standards are supplied and the test is far more comprehensive, thorough and reliable. There are after all only two distinct types of arithmetic and these must be tested somewhat-as our fathers tested them though we have at our disposal a more thorough-going way of doing this.

The two types of arithmetic are computation and practical problems. There are many processes to be taught under computation, some of them quite complex, but they have this in common. they are strictly mechanical and objective. There are certain rules which pertain, and if we but apply them rightly, the answer must be one and the same for anyone who works them. Problem-solving on the other hand cannot be reduced to a mere mechanical application of rules. Problem-solving is a matter of first understanding the situation (each problem that is not mere repetition with only a slight change in data presents a new situation). After the problem is properly understood (and this requires intelligence) the data must be properly arranged and it must be decided what manipulation of this data will give the information called for. Thirdly, accurate mechanical computation is required to work out this information. It is not at all infrequent that the two abilities, namely, solving problems and doing computation, are not found in the same person. The one is almost exclusively a mechanical skill and the other is common sense plus good language ability plus skill in computation.

Problem-solving is to such an extent a natural outcome of 'the common sense we are born with' that there are many people whose ability to solve problems is not greatly helped or hindered no matter what the teacher does. This is not nearly so true in the case of computation which is one of the most teachable subjects in the whole school curriculum. A pupil who is good in language but weak in computation is farther along the road to problem-solving than the one who is good in computation but weak in language. The computation defect can be more easily remedied than the language defect. Computation takes much practice but problem-solving takes good natural ability in addition to practice.

Many inspecting officers today almost exclusively give practical problems in the examinations they set. This is unfair to the less gifted pupil and the teacher, for problem-solving is so much a matter of natural common sense which we either have or do not have, regardless of how hard we work and how hard the teacher works, that these officers might almost as well give an Intelligence test and judge arithmetic by that. Pupils who have not enough natural ability to understand the why and wherefore of the problems, can be taught to do a few forms of hard practical problems, but they will soon forget the process and muddle it up hopelessly with that of some other problem that they learned equally mechanically and unintelligently. Much of the preparation of students for examinations is of this type, forced upon teacher and pupil by the unreasonably hard problems given in examinations.

In Standardized tests, problem-solving and computation are usually quite separate, for they are considered as different from each other as English and Geography. The part of the test dealing with practical problems is simply a sample selection of various types of problems. Some of these will be rather stereotyped and some are particularly designed to tax ingenuity and common sense. Sometimes an entire section of a test will be given on arithmetical reasoning, in which a large number of short problems are given which demand thinking more than computational skill or familiarity with stereotyped procedures. Much more cleverness is required to do a large number of short easy problems which demand a variety of solutions than to do long complicated problems of traditional types. Some tests give as many as forty problems to be solved in as many minutes.

In computation one differentiates between tests of *speed* and tests of *power*. The power test tries to find out how

difficult a piece of work the pupil can do. The speed test tries to find out how much work of a certain difficulty a pupil can do accurately in a stated time. Most tests contain both speed and power exercises.

Speed is nowadays considered much more important than formerly, not only because a speedy worker saves time but because speed is an indication also of the ease and efficiency with which the work is done. Someone who works hard while playing the piano will not play with the certainty and mastery that another exhibits who feels relaxed and at ease while playing. Thus it is in skills; we become skilful in doing a certain intricate piece of work when it becomes second-nature. Speed indicates freedom from uncertainty and defective habits of work, and a degree of mechanical mastery that will not quickly wear off. We may assume that he who can do correctly a large number of computations in a short time does his computation with commendable workmanship and will not soon lose his skill.

Educators have found that it pays to demand high standards of speed and accuracy in simple computations because the efficiency gained in easier work prepares the pupil for harder work; and because harder computations are really not required in practical life nearly as much as the simpler forms. Only occasionally does the non-expert have to make calculations well up into thousands. We seldom use long calculations but what calculations we do need to make, we need to make accurately. The use of printed practice material has given pupils many times the practice opportunity they formerly got. Where printed material is not supplied, much time is spent in tasks like copying rather than in calculation. This emphasis on accuracy and ease in workmanship stands in contrast to the practice in India where as low as one-third accuracy

is tolerated and time enough allowed for any number of trial-and-error starts.

A specimen page from *The Northumberland Arithmetic Test* by Cyril Butt¹ is herewith given. This is an exercise in addition

Add these figures as quickly and carefully as you can:

		s.	d.	ft.	in.
12	53	8	2	2	4
34	49	3	1½	3	1
10	75	2	5	5	3
<u>28</u>	<u>62</u>	<u>6</u>	<u>3</u>	<u>4</u>	<u>2</u>

(Do not reduce feet to yards)

lb.	oz.		493	£	s.	d.
2	5	47	502	4	7	11½
1	4	82	768	3	16	4½
6	3	59	916	5	2	5¾
<u>7</u>	<u>2</u>	<u>16</u>	<u>857</u>	<u>8</u>	<u>9</u>	<u>1</u>

yds.	ft.	in.	tons	cwt.	qr.		
4	2	9½	5	17	2	5½	89.086
2	1	7½	3	5	3	3½	17.321
5	0	8½	3	12	1	8½	0.965
<u>6</u>	<u>2</u>	<u>5½</u>	<u>7</u>	<u>16</u>	<u>0</u>	<u>2½</u>	<u>24.403</u>

Add these figures UP and ACROSS, finding the totals for every year and every town.

	<u>1920</u>	<u>1921</u>	<u>1922</u>	<u>1923</u>	Totals
London	9,146	8,639	7,284	6,789
Manchester	5,295	6,472	5,973	5,531
Liverpool	6,374	5,837	4,625	4,247
Newcastle	<u>4,087</u>	<u>4,985</u>	<u>5,167</u>	<u>3,556</u>
Totals
			Grand Total	

¹ By kind permission of the University of London Press.

There are other pages on subtraction, multiplication, division, mental arithmetic, rules, and arithmetical reasoning.

Of a different type is the *Monroe Survey Test*.¹ This seeks to find out the speed of working examples of a specified uniform quality. This is distinctly a speed test. The pupil is stopped after four minutes of work. Fourteen such examples are given in addition alone. Similar exercises follow in multiplication, division, subtraction, simple fractions, and the placing of the decimal point in division.

7682	6809	8941	5917	6772
8975	7883	6028	4814	7910
9005	8240	6535	9007	9845
5872	6601	8522	6975	2340
<u>3203</u>	<u>3739</u>	<u>6794</u>	<u>3426</u>	<u>2319</u>
7917	3293	6794	8758	2462
4304	7805	5420	2350	9869
9027	7642	3197	4572	1081
2838	9975	7803	8028	4314
<u>1247</u>	<u>2462</u>	<u>8758</u>	<u>1249</u>	<u>5917</u>

As an example of a power test observe the exercise following this paragraph. This test starts with the easiest work in simple division and continues through step after step of increasing complexity and difficulty. The object of this test is to find out just how far in simple division the pupil can go before he strikes the stage at which his work is uncertain. Such a test is of value to the teacher in finding out where to begin to improve work in division. Liberal time is given in this test because the information wanted is not in regard to speed but to difficulty.

¹ By W. S. Monroe. Published by the Public School Publishing Company.

Series 1.	3) 9	7) 49	6) 54
„ 2.	4) 8448	7) 1477	6) 1266
„ 3.	5) 2055	4) 237	9) 4185
„ 4.	12) 288	17) 5899	15) 4035
„ 5.	27) 432	36) 17408	73) 2475
„ 6.	42) 521793	58) 300309	94) 594321
„ 7.	324) 357449	648) 3648081	487) 3387401

The following arithmetic tests, standardized for India, have come to the writer's notice.

1. *On the primary level* there is one in Kanarese by D. V. Chichermane of Belgaum. He examines addition, subtraction, multiplication, division, and arithmetic reasoning under separate sub-divisions including examples from the easiest to the most difficult to be expected in primary schools. There are age norms as well as class norms.

There is also a Hindi and Urdu test by the author which goes through the work included in the syllabus of each class of the primary school separately. Every examinee is expected to start at the beginning and work on until the examples are too hard or the process is unknown. In this way we not only have a check on accuracy but can easily find out the level which the pupil has reached. If a fourth-class pupil is weak not only in the work of Class IV but in the work of Class III and even Class II, one knows immediately how far back revision should be done. In this way such significant facts were found out as that the majority of fourth-class pupils do third-class work so poorly as to suggest that entirely too much is crowded

into the primary school course. Third-class pupils are just beginning to get an acceptable grasp of the work of Class II which makes it not at all surprising that the standard in the work prescribed for Class III is so low. This is distinctly a power test which tests the appropriateness of our present curriculum as well as the ability of the individual. Testing on a wide scale tests the curriculum and the test as well as the pupil.

2. *On the middle school level* there is a single general survey test (a general survey test is one which makes a rapid survey of the principal processes in the subject as a whole. In other words it is a general review test) in both computation and problems. This forms two parts of the Middle School Achievement Test (Hindi).¹

3. *On the high school level* there is a very good general survey test in both computation and reasoning by N. C. Chatterjee of Patna (in English, Hindi and Urdu). The section on computation covers a very wide field and is quite thorough as a survey. Copies are to be had from the Patna Training College.²

For both the middle and high school level there is also a very handy scale of arithmetic problems worked out by Mr S. C. Chakravarty of the David Hare Training College in Calcutta.³ This scale was originally used in the Bengali language. It appears in English in *Teaching* but its trans-

¹ 'A High School Entrance Test in 2-4 Hours' (in nine subjects), *Teaching*, Vol. XIII, No. 2, December 1940. Oxford University Press.

² 'A Standardized Objective Test in Reasoning Arithmetic' by N. C. Chatterjee, *Education*, March 1940 and 'A Test in Mechanical Arithmetic' by the same author in the *Indian Journal of Education*, September 1939.

³ 'Reasoning Test in Arithmetic', by S. C. Chakravarty. *Teaching*, Vol. X, No. 1, September 1937. Oxford University Press.

ation into any language should not cause much difficulty nor a very appreciable difference in the norms if the translator tries to make the translation as simple as possible.

There are twenty problems in the scale carefully graded in difficulty by fairly regular gradations. With the kind permission of the author, the scale is printed entire

REASONING TEST IN ARITHMETIC

(Time allowed: 45 minutes)

1. There are 105 mangoes on a tree, of which 17 are ripe. If the wind blows down the ripe mangoes, how many are left on the tree?

2. 300 mangoes are put into 5 baskets so that each basket has the same number of mangoes. How many are there in each basket?

3. Two groups of boys are standing apart. In the first group there are 77 boys. If the two groups stand together, the number of boys becomes 183. What is the number of boys in the second group?

4. Ram Babu wants to buy a house for Rs. 5,000. He has Rs. 2,340 in his pocket. How much more money does he require?

5. After giving 5 marbles to each of 27 boys, I have 13 left. How many marbles had I at first?

6. By selling a watch for Rs. 115 there is a loss of Rs. 17. What is the real price of the watch?

7. Fifteen seers of water can run out in a minute through a pipe attached to a cistern. If the cistern can be emptied through the pipe in 12 minutes, how many seers of water does it contain?

8. If a cart is sold for Rs. 700, there is a loss of Rs. 125. For what should it be sold to gain Rs. 100?

9. The product of three numbers is 3,640. If two of them are 13 and 14, what is the third number?

10. On spending Rs. 2,527 in a year a man incurred a debt of Rs. 470. How much could he have saved if he had spent only Rs. 1,735?

11. After dividing 200 oranges equally among his 38 classmates and himself, a boy gave what was left over to the bearer of the school. How many oranges did the bearer get?

12. Satish has as many rupees more than Bimal as Ramesh has less. If Bimal and Ramesh have Rs. 125 and Rs. 90 respectively, how much has Satish?

13. A person has Rs. 186 in one box and Rs. 62 less in another. If Rs. 13 is transferred from the first box to the second, how much more will one box contain than the other?

14. A person gives Rs. 7 and Rs. 11 to the poor on every Wednesday and Saturday respectively. In what time will he give Rs. 576 to the poor?

15. Seven years back the sum of the ages of Bimal, Satish and Ramesh was 34 years. What will be the sum of their ages two years hence?

16. A dealer in fruit, after selling some oranges to 18 persons found that he had still two rupees' worth of oranges left. If he had sold two more oranges to each of them, he would have had only 4 oranges left. How many oranges did he sell for a rupee?

17. There are 1,652 mangoes on three trees. The second tree has 340 mangoes less than the first, and the third has 112 mangoes more than the second. Find the number of mangoes on each tree.

18. Some birds are perched on two trees. There are 17 more birds on the first tree than on the second. If 5 birds go over from the second tree to the first, how many more birds will one have than the other?

19. The price of a horse and carriage is Rs. 1,200. If the price of the carriage is Rs. 20 less than four times the price of the horse, what is the price of the horse?

20. A person saved some money by spending Rs. 600 annually for 5 years. But on spending Rs. 720 annually during the next 7 years he incurred a debt of Rs. 480. What was his annual income?

The norms for this test, as given by Mr Chakravarty are:

Class	Problems	Right
V	6.5	
VI	7.4	
VII	9.1	
VIII	10.9	
IX	13.1	
X	14.1	

Every teacher of mathematics in India should administer either this scale or such scales in arithmetic problems as are found in the chapters on mathematics in *Mental and Scholastic Tests* by Burt or in *Mental Tests* by Ballard (see bibliography). The use of such a scale together with a careful analysis of the results will lead to a much clearer conception on the part of the teacher of the degrees of difficulty of arithmetic problems which one can fairly expect pupils of various ages and accomplishments to solve. The use of such a test will give results which most teachers will regard as very unexpected indeed. Few teachers will escape having to revise their ideas of what difficulty of problems are appropriate to each class.

There is real need for Standardized Arithmetic tests in India. If these were available with norms, the urgent question of whether or not the curriculum is at fault for the large number of failures in arithmetic could be settled on the basis of fact instead of mere opinion. Evidence thus far accumulated seems to indicate that the arithmetic course is very poorly and unjustly adapted to the ages of the pupils of each class.

An article which appeared in *Teaching*¹ on the 'Arithmetical Abilities of Indian and American Students Compared' may prove of interest to the reader.

¹ Vol. IX, No. 1, September 1936.

LANGUAGE TESTS

A. The Mother-Tongue

In India the mother-tongue in all but a small proportion of schools will be one of the Indian languages. English is the mother-tongue of but comparatively few in India.

There are probably no fully Standardized Language tests in any Indian language. If there are any, they have not been brought to the attention of the writer.

It must be borne in mind that the language treated as the mother-tongue in schools is very often not the actual mother-tongue of many of the pupils. Wherever a dialect flourishes, the pupils have to learn the official school language almost like a new language. In some cases it actually is a new language. Some pupils are bound to be handicapped, especially since outside the school building they rarely speak the official school language. Not enough consideration is given to this fact. It is unfair to demand the same mastery (especially in the case of primary school pupils) of an official language in an area where a dialect is in universal use, as is demanded of those whose home language and school language are exactly the same. When some good Language tests have been standardized we shall have the means of finding out the extent of the hardship imposed on those living in dialect areas. Since the dialects are not taught in school, the term 'mother-tongue' is here used in the sense of official language, whether this is the actual mother-tongue or not.

Reading comes under the head of language studies but this topic has been treated separately because of the peculiarity of its problem for testing. A good many kinds of work are included under the heading of language including (besides reading) grammar, word studies, spelling, composition, literature, and rhetoric.

The Sones-Harry High School Achievement Test,¹ which is considered one of the very best in its field, includes the following sub-headings under the general heading 'Language'.

1. *Correct and faulty use of the language.* This is a True-False test requiring that the sentences containing faults be marked wrong and those without faults right.

Examples

They left the neighbours and I at home.

He would carry neither life nor accident insurance.

2. *Word Meaning.* The directions are to underline the word in brackets that means the same as the word in italics.

Examples

specific (general, definite, regal, unitary)

frankly (candidly, slowly, roughly, convincingly)

3. *Abbreviations and prefixes.* This is a Matching test in which abbreviations such as A.D., i.e., pro tem and prefixes like ad-, ante-, post-, are given in one column and the meanings in another. The expression and the meaning must be correctly matched.

4. *Why certain common expressions are faulty.* This exercise is also a Matching test containing examples in which italicized expressions are wrongly used. In another column are given a series of rules which have been violated. The pupil has to match the faulty expression with the rule it violates.

5. *Foreign phrases used in English.* Such foreign expressions as: 'dramatis personæ', 'tête-a-tête', and 'nom de plume' have to be correctly matched with their meanings.

¹ By W. W. Sones and D. P. Harry. Published by the World Book Company.

6. *Literary forms* Again a Matching test giving in one column various forms of literature such as ballad, comedy, elegy, epic, etc. and in the other the names of well known literary products such as *Ivanhoe*, *The Iliad*, etc

7 *Reading comprehension*

8. *Authorship of famous literary products of various nations*. This is a three-column Matching test. In one column are the names of the authors. These must be matched with some item in the column containing the names of the literary masterpieces. The nationality of the author must also be matched.

9. *Familiar characters of literature*. In one column are brief descriptions of famous characters of literature to be matched with the names of the characters in the other column.

10. *Familiar passages in literature*. Poetry passages are given from which one line has been omitted. In the other column are missing lines which have to be put into the correct passage. The poetry passages are from masterpieces with which the examinee is expected to be familiar.

11. *Literary themes*. The themes of certain well-known books, are to be matched with the titles of the books which treat the various themes.

12. *Technical vocabulary of language*. Such terms as: clause, idiom, phrase, etc., are to be matched with the proper definitions. There are other series of items on the identification of grammatical and rhetorical forms, characters famous in literature, American and English authors, and the literary interests of various authors.

This is a comprehensive test. It is impossible to imagine any test of the traditional type covering such a broad field in the brief space of the forty minutes allowed.

*The Metropolitan Achievement Test*¹ for Classes IV, V and VI contains a section on reading, one on vocabulary (multiple-choice), and a section on punctuation and capitalization in which sentences are printed without capitals and punctuation, to be corrected by the pupil. There is also a section on spelling.

The above tests have been outlined in order to give an idea of what may be included in such tests. Language tests approach their work in a systematic way, using lists of the more common faults and difficulties in order that the tests may really bear on the strategic points of good usage.

There are also composition scales to be had. In these scales short compositions are listed in descending order of merit from the very good to no composition value at all. Although one of the hardest fields in which to get objective measurement, standardized evaluation of composition was one of the first problems attacked in the whole educational measurement movement. Milo B. Hillegas published his *Scale for the Measurement of Quality in English Composition for Young People*² as long ago as 1912, when intelligence testing was just beginning to claim a little public attention. It consists of ten compositions selected from 7,000 samples graded in order of merit by over 200 qualified judges. *The Hudelson Typical Composition Ability Scale*³ is claimed to be the most generally useful. There are a large number of similar scales in the market.

¹ Edited by J. S. Orleans and published by the World Book Co.

² Not in general use nowadays. Remodelled into *The Hudelson English Compositions Scale* and published by the World Book Company.

³ By E. Hudelson and published by the Public School Publishing Company.

It takes some skill and practice to be able to use such a scale correctly. Nevertheless, the study of one is extremely valuable to a teacher in order to keep before him gradations in order of merit. Even if the scale is not used regularly, it sharpens his sense of evaluation and aids in familiarizing the teacher with what is to be expected of pupils of various ages and school grades.

Word-counts in various languages

Word-counts are not tests but have been so intimately connected with the development of standardization and careful grading of school work according to the capacity of the pupils of certain ages and school grades, that the matter should be mentioned here, especially since some valuable work has been done in this respect in India. Most standardized tests in Language use word lists prepared from word-counts.

A word-count is simplicity itself. The reason for a word-count is that it is pedagogically incorrect to burden young children with many unfamiliar words at one time. We should let pupils read and hear language which they understand thoroughly so that reading may always be meaningful to them, introducing new words slowly but systematically. The words one can expect everyone to know are obviously those which are most used. If we list the most common words we also list the most familiar words. Instead of merely guessing as to which the common words are, the more accurate way is to count the times various words are used.

In the English language ten million running words in books, newspapers, magazines, and other printed matter

were counted, giving us the famous Thorndike list¹ of the 15,000 most common words of the English language.

Since the publication of this list, reading books have been constructed on an altogether different basis. In America and England no respected educationist would set about preparing 'readers' for young children without consulting such a word list as Thorndike's in order that he might begin with the more familiar and hence, generally speaking, easier words.

In India such word lists in all languages are even more necessary than in English in England or America, for the systematic use of the most familiar words would tend to make them even more familiar and decrease the evil of India's population being split into so many different tongues. Especially where a dialect is in use, the restriction of vocabulary is very necessary since under such circumstances the school language is not the mother-tongue but almost a foreign tongue. Any person who has to learn a foreign language will remember how his first use of the language was confined to the few words he knew. Unless those who talked to him restricted their vocabulary he could not understand. The learning process was speeded up if a few new words were gradually introduced but when too many were introduced at once he could no longer understand and the learning process stopped with his feeling simply helpless. A word list permits the introduction of new words to be made systematically according to the capacity of the learner.

Mr J. C. Koenig made the first scientifically constructed word list in an Indian language. It is in Hindi. The first

¹ Thorndike, E. L., *Teachers' Word Book of Twenty Thousand Words Found Most Widely in General Literature*. Bureau of Publications, Teachers' College, Columbia University, 1922.

list contained only 1,000 words. Later the Department of Education of the Central Provinces co-operated and the result is a list of the 4,000 most common words in Hindi.

This word list should be in the hands of every editor and writer who wishes to teach in Hindi others besides the highly educated. Of those considered literate in India, only two per cent have finished high school, which means that if one would reach more than a handful of India's great population one should so write that even the primary school boy or girl can read what one writes. Hindi has a tremendous vocabulary, of which people of little education can be expected to know but a small part. If writers would confine themselves to the most familiar words of the language, they would sound less highly educated but they would certainly make themselves better understood. The function of language is to give one a medium for making oneself understood. Let writers confine themselves to a more basic language and the present complaint that so few books, papers, and magazines are sold may be partly removed as the less highly educated find that literature is more readable than they had found it to be in the past.

The many advantages of such a word list, as well as the method of preparing them, are concisely but fully described in the English introduction to a separate edition of the word list which is published by the Mission Press at Jubbulpore. This introduction will also be suggestive to those interested in word lists in other languages than Hindi.

This Hindi list is also to be found in the Hindi edition of the *Handbook for Teachers* prepared by the Department of Education of the Central Provinces and published by the Indian Press, Jubbulpore. A similar list in Marathi

to be found in the Marathi edition and one in Urdu in the Urdu edition of the same work.

Within a few years, as the adult education movement sets under way, there will probably be similar lists in other Indian languages.¹

4. Foreign Languages

As far as I am concerned, all languages are divided into two great classes, the languages which I spoke in my childhood and the languages I did not speak in my childhood, or, my mother-tongue or tongues, and those that are not my mother-tongue. In between these another class of languages has wedged itself, namely those with which I have acquired some familiarity but in which I simply cannot feel as much at home as in the two languages I spoke as a child. One of these later-acquired languages I have used for twelve years but it still is in a decidedly different class from my mother-tongues.

A foreign language is approached differently by us from our mother-tongue and this difference will be reflected in both the teaching of the language and the demands a test in it makes.

The method used in teaching will do much to determine the kind of test that should be prepared. The kind of test prepared, again, will probably influence the methods of teaching employed in subsequent years.

¹ There are, at the time of writing of the second edition of this book, word lists in practically all of the most commonly spoken vernaculars of peninsula India as well as in some dialects. Some of these are based on language as found in literature (constructed on the principle of the Koenig word list) but some of these are based on language used in conversation. One should always inquire how a list was compiled in order to avoid adopting wrong inferences from the list.

Teachers of Sanskrit and Persian may find standardized tests like V. A. C. Henmon's or D. S. White's *Latent Tests*¹ suggestive.

In India the foreign language which is of most importance is English. Because of the numerous opportunities and necessities for both reading and speaking English, this language has to an Indian student a less foreign flavour than French or German. Since English has been used as the medium for teaching a number of subjects it is actually easier for some Indians to converse about these subjects in English than in their mother-tongue. Take the subject of psychology, for instance. Probably few Indian psychologists would be prepared to offer a scientific paper on this special subject in their mother-tongue without worrying about it a good deal. No matter how great the demand for 'vernacularization' may become, there will still be a demand for English to be taught as a spoken language. The direct method is an attempt to teach a language as a child learns it, emphasizing hearing and speaking rather than studying from a book. This is in contrast to the method which immediately plunges into grammar and the memorizing of long vocabulary lists. The one method encourages us in one way or another to express our ideas and wants in the new language, no matter how many mistakes we make. We are to learn the language, by speaking it. The other method inhibits us from expressing our ideas and wants, until we have mastered the grammar of the new language. Consequently we keep our ideas and wants to ourselves which results in retardation in acquiring the new language and in a bookish language when finally we do learn to speak. Since English is wanted for use in everyday life more

¹ World Book Company.

than it is wanted for reading Shakespeare and Shaw, it should be taught as conversationally as possible.

Pronunciation can be tested only by the one who hears the language pronounced. This important element of learning a language does not lend itself to standardization on a wide scale.

Mr George Leeder, the principal of Ewing High School in Ludhiana, has adapted a reading test which was intended for children in Classes III and IV whose mother-tongue is English, for use in middle schools in India among pupils who learn English as a foreign language. This test lends itself to use quite readily where the direct method has been used. Thorndike's word list was used in composing and re-adapting it so that the vocabulary demands are just about what one might expect of one who tries to work his way into the language through conversation. Especially the part on the reading of simple paragraphs should encourage emphasis on getting the main thought of the speaker (or writer) instead of upon details of vocabulary and sentence construction (which emphasis is characteristic of the teaching of dead languages rather than living).

One of the paragraphs of the test is given here for illustration:—

One day we had a post office at our house.

John played postman.

Sundar worked at the post office.

We wrote letters to Jane and Yusuf and put them in a toy post box.

The postman took the letters to the post office

Sundar sent the letters to our friends at school.

(a) We played post office

at school in the bazaar at home

(b) A letter was sent to

John Sundar June Rama

(Underline the correct word or words)

An English child in Class III who reads this paragraph does an entirely different thing from what an Indian child in Class VI does. For the English child the test is not a matter of understanding the contents of the paragraph. He would understand it perfectly were the paragraph spoken instead of being printed. His problem is to get the print to have the same meaning for him as the hearing of a story has. The Indian child, on the other hand, is assumed to be a better reader in his own language than the younger English child and does not find the reading of the English letters of the alphabet particularly disconcerting, for if he would practice reading his own language in Roman script he could soon do that as well as reading the Devanagari. But what does matter to him is that the words are not his mother-tongue but a foreign language. His task is to understand the language of the paragraph whether it is given to him in print or orally. Orally it would probably be harder for him, which is opposite to the case of the English child whose difficulties are assumed to be in the mechanics of reading rather than in the understanding of the language. This is not a *reading test* for the Indian pupil but a *language test*. His ability to understand a foreign language is being tested. English reading tests should be suggestive in constructing not only language *tests* but language *exercises* for Indian students. However, these 'Reading tests' should be used with caution. The contents of English books and paragraphs chosen for Reading tests assume a background that is foreign to the experience of the Indian child. Many of the experiences of an English child, even

though they are told in his mother-tongue, are so strange to the Indian child that he will not understand them readily. The contents of all material must be thoroughly adapted, taking into consideration the different background of the Indian child.

Part II of Mr Leeder's adaptation is a Multiple Choice test on the meaning of words.

A hat is for the (foot, hand, face, head).

A loud noise is made by (butter, bungalow, thunder, autumn).

The final part of the examination is a Vocabulary test (choice of the proper synonym from among several given words).

English is the easiest of all subjects in which to develop Standardized tests, for one and the same English test is usable all over India, and is not restricted by the barriers between the various languages.

The test outlined above does not 'measure' language expression on the part of the pupil. The usefulness of a brief English composition scale for Indian students might be explored. Any scale made for the use of English-speaking children will not be suitable. The scale should be worked out anew for Indian conditions and reflect the background of the student who is seeking expression in a tongue other than his mother-tongue. It might be that college students in India could use one of the scales current in England or America, but lower class pupils can hardly be expected to find English a natural enough means of expression to make their composition difficulties at all commensurate with those of persons who were born in English-speaking homes.

Mr Rahim, Headmaster of the Anjuman High School in Jubbulpore, has worked out some quite usable English tests based on the vocabulary of the textbooks used. These

include some really fine types of vocabulary exercises along the lines often used by the newer English textbooks. Simple paragraph work is also included.

Many English intelligence tests include items which test very well the ability of students to understand English. Such items require not only that a student is intelligent enough to carry out the instructions in the intelligence test but enough mastery of English to understand what is wanted. When using intelligence tests in this way, items which would be easy for the age group if given in the mother-tongue should be used.

Before a good test in language-usage difficulties can be composed for Indian students someone will have to record and classify the mistakes which Indian students are most likely to make. The habits and practices of our own mother-tongue tend to carry themselves over into the foreign languages we are learning, especially if we do not 'think' in the foreign language but translate into it from our mother-tongue. Indian idioms are bound to superimpose themselves into English in the case of Indian students, just as the Hindi spoken by Englishmen and Americans turns out to be a peculiar 'Sahib Hindi' which causes much amusement to those whose mother-tongue is Hindi. 'Sahib Hindi' and 'Babu English' are twin linguistic phenomena.

During my first schooldays in America, the teacher often had to remind me not to say 'I ain't got no pencil' when I meant 'I haven't any pencil'. An Indian learning English will not be tempted to use this form of faulty expression but will probably have to be reminded that 'He is too honest' does not mean 'He is very honest'.

A thoroughly standardized test for India will not content itself with dealing with the language difficulties of those who speak English as a mother-tongue but will

study the peculiar mistakes in grammar and idiom to which Indian initiates in the language are liable.

Where an Indian language is taught as a second language (like Hindi in Gujarat), there is no reason why a Standardized test prepared for the Hindi area should not be given here also. But the norms for the Hindi area cannot be held to be a fair standard for any area where Hindi is not commonly spoken or at least not spoken commonly by the examinee. One could not expect students of Hindi in Porbander to compete with those in Jubbulpore. The standard demanded would have to make allowance for the local handicaps. On the other hand students in Lucknow studying Hindi might be expected to be not nearly as much handicapped as the Porbander students, even though their mother-tongue is Urdu, for Urdu and Hindi have more common elements than Gujarati and Hindi.

Where to get samples of School-subject tests

There are many other school subjects for which excellent Standardized tests have been prepared for use in America and England such as hygiene, geography, history, physics, chemistry, algebra, and shorthand. It will not be feasible to discuss these at length, for there is hardly space in this book to give but the briefest description of any type of test. Nearly all of the above subjects are informational in nature and all tests of largely informational subjects can be constructed along somewhat similar lines, as illustrated in the True-False, Multiple Choice, Matching and other tests. There is no essentially different principle of construction to be borne in mind as in the case of reading or writing that is not already illustrated in the treatment of the differences between the old-fashioned type of written examination and the new type.

No one can get an adequate idea of what New-Type Standardized school tests are until he has examined a good number of these tests and attempted to give at least a few of them. He should also occasionally be an examinee in such a test. There are few better ways of learning humility and understanding of the pupil's position.

Principals, teachers, and inspectors in India have not found it convenient to obtain such tests. There are very few available composed expressly for India.

Practically no imported test is usable in any but a limited way exactly in the form it comes to us. Nevertheless such tests contain such a wealth of teaching suggestions that no progressive teacher should neglect to have at least a few tests on hand in the subjects he teaches. There is probably no literature that can give more suggestive illustrations for the improvement of teaching methods and aims. It would take the reading of many books on the teaching of mathematics, for instance, to give us as vivid an impression of the changed outlook, standard, and aims to be stressed in the teaching of arithmetic and geometry as the study of a few recognized tests can give. Most of these tests have been composed by the outstanding pedagogical experts of their field.

Such tests should by no means be used in the way that books of the *Questions-of-the-Matriculation-Examinations-of-the-Last-Fifty-Years-Fully-Answered* type are used. Standardized tests should never be used for coaching purposes. But they might be used by the teacher to keep sounder pedagogical objectives in mind as he teaches, and to offer suggestive material in the preparation of his own informal tests which he prepares for teaching and diagnostic purposes as well as for examining purposes.

All the books by Ballard and Burt mentioned in the Bibliography contain numerous samples of School-subject

and Intelligence tests. These books contain the most readily obtained samples of tests for the Indian teacher. Both Burt and Ballard are considered among the most able of authors of New-Type tests. American tests are not obtainable in book form. The development and use of New-Type tests has gone much farther on the other side of the Atlantic so one can hardly get an adequate idea of what New-Type tests are without importing some from the United States.

A list is given of some of the best known publishers of test material who are in a position to supply a large variety of samples. Each of them stocks well-known tests. They will send their catalogues on application.

Some years ago the author purchased a small filing case of test material from the World Book Company entitled *A Complete Outfit of Test Material*. The price was ten dollars (about Rs. 27-8). There are probably about two hundred tests of all varieties from pre-school to college, including School-subject tests, Achievement tests, Mechanical Aptitude tests, Intelligence tests, Personality and Attitude scales, etc. This file is a mine of information. The price quoted above would naturally be enhanced by postage charges. Such a file of samples should be in the staff library of every progressive high school or college. Every staff member would find something suggestive and helpful in it.

An individual would not need such a large variety of samples and might not care to incur such a considerable expense. Any of the publishers listed would send a few representative samples in special fields. A mathematics teacher might send ten shillings by postal order and ask for as many samples of tests in arithmetic and higher mathematics as can be given for the price enclosed. Or a teacher of science and language in a middle school

might ask for samples of material in these two subjects for pupils in Classes V to VIII. The publishers would undoubtedly try to be as helpful as possible.

The Public School Publishing Co. advertises the following sample packages and folios of tests:

Sample Package 'A'. Tests for elementary school subjects, \$ 1.50, post paid.

Sample Package 'B' Tests for kindergarten and primary grades, \$ 1.50, post paid.

Sample Package 'C' Tests for high school subjects, \$ 2.00, post paid.

These sample packages contain complete sample sets of representative tests in the subjects indicated.

Portfolio of Tests for use in Teacher-training Classes including 18 sample sets of tests, \$ 1.50, post paid

Folio of Reading Tests, \$ 1.50, post paid.

Folio of Arithmetic Tests, \$ 1.50, post paid.

Folio of English and Composition Tests, \$ 1.00, post paid.

Folio of Mathematics Tests (Algebra, Geometry, and Commercial Arithmetic), \$ 1.00, post paid.

Folio of Tests in Social Studies, \$ 1.00, post paid

Folio of Science Tests, \$ 1.50, post paid.

Folio of Latin and Modern Language Tests, \$ 1.50, post paid

Each folio of tests is composed of a collection of complete sample sets of tests in the particular subject, for both elementary grades and high school, where available

More publishers could be listed but the following will probably give a wide enough selection:

1. Board of Education, London, England.
2. Bureau of Publications, Teachers' College, Columbia University, New York City.
3. The Educational Test Bureau, University and 15th Avenue, Minneapolis, Minn., U.S.A.
4. C. A. Gregory Co., 345 Calhoun Street, Cincinnati, Ohio, U.S.A.

5. The London University Press, 10 Warwick Lane, London, E.C 4.
6. Public School Publishing Co , Bloomington, Illinois, U S.A.
7. World Book Co , Yonkers-on-Hudson, New York, U S.A.

The preponderance of American publishers in the list is due not only to the author's national allegiance but to the fact that the great majority of Standardized tests available are of American origin.

Little effort has been made to list the names of the better-known tests in various fields, since the resident in India will not find it convenient to write to several publishers for samples but will prefer getting them from the same publisher, which will suit his purpose just as well. Each publisher usually handles only his own material, but the C. A. Gregory Co. handles the publications of various publishers.

IV

ADVANTAGES AND LIMITATIONS OF NEW-TYPE EXAMINATIONS

SINCE we have now finished the discussion of school-subject examinations, and the types of examinations to be described in the succeeding chapters touch entirely different problems, it will be well to sum up the question of in what way New-Type tests can influence examining methods in the school.

Much of the contents of this chapter has been implied in previous pages. In describing the New-Type tests, an attempt has been made to show something of their advantages and limitations; but it is necessary now to bring this information together in organized and concise form. The full explanation of the advantages and disadvantages of New-Type tests was not considered desirable until it had been shown what New-Type tests are.

Percival Symonds in his *Measurement in Secondary Education*¹ gives what he considers the five advantages. These will be commented upon individually.

First Advantage

✓ The new-type examinations are more objective in their scoring, since the responses of the pupils are controlled, and since there can be no doubt as to the correctness or incorrectness of their responses.

As to the objectiveness of the scoring, it has already been mentioned that recently a machine has been invented

¹ Pages 39 and 40.

which can do the scoring. Any test that is so worded and arranged that a machine can do the evaluating, can certainly not be accused of being dependent on the varying idiosyncrasies of different examiners. There may be argument as to whether the machine should have been set to score a 'No' or a 'Yes' on question 56, but there can be no question of the machine marking the same answer right on one paper and wrong on another unless it is mechanically faulty. Such machines are very new and expensive and one cannot expect the evaluating of tests in the average school to be done by machines. And yet the human being as he marks papers of the new-type tests is much more machine-like than when marking papers of the old type. He may make occasional careless mistakes or even get mixed up in the use of his scoring key but this does not happen often and can be guarded against by re-checking.

Compared with this, the evaluation of the traditional essay-type examination is extremely subjective and variable. Take a legitimate thought-provoking assignment like 'Contrast the statesmanship of Akbar with that of Aurangzeb'. This is a splendid theme for a student of history to work on, but it does not make a good examination question on which to decide the fate of thousands of students, because when it comes to evaluating the paper, what one examiner considers a good answer may be quite different from what a second examiner calls a good answer. If each of one hundred papers were examined by two different examiners, it is probable that at least 25 per cent of the students would be failed by the one examiner and passed by the other. This assertion is not based on guess-work but on investigations such as those reported in *An Examination of Examinations*. In this book it is reported that in the evaluation of the

English essay of a college entrance examination, the average difference between the marks assigned by several examiners to the same paper was 19.6 out of 100. On some of the papers two examiners differed by as much as 37 out of 100. In the papers of some other subjects the differences on the same paper were as high as 52 out of 100.

In India there is much discussion as to whether high marks in one paper may be permitted to compensate for a shortage of one or two marks in another paper. These discussions sound like exact mathematics, but they begin to look foolish when we realize that in most cases another examiner who does not know what marks the first examiner assigned, in going over the same paper would quite possibly score it twenty marks higher or lower, it being a toss-up which. If the same examiner were to get the same paper a second time without seeing the notes on how he had evaluated the paper before, the difference might be equally great. The fine distinction attempted in India between 'second' and 'third' division is none too reliable when it is realized that the differences in marks assigned by different examiners, or by the same examiner on different dates, average almost as much as the difference in the average marks of these two divisions.

Mr Kuppuswamy Aiyangar reported just as great an unreliability in India¹ as the aforementioned report on examinations in England

Second Advantage

They may be very comprehensive and can be made to cover a great deal more material than the old type of examination. Since the pupil does very little writing,

¹ *The Present Examination System*, printed at the V. V. Press, Trivandrum, Travancore, 1931.

he can devote his time to thought and can thereby answer a great many more questions than when he has to write his answers out at length

A short test is always unfair if any but a very rough classification is aimed at. If the questions be of a broad nature and deal with only the fundamentals of the entire course, an examination of but a few questions need not necessarily be unfair because of the small number of questions. What the test lacks in the number of questions it may make up in comprehensiveness. But where details of fact or accuracy in a skill like mathematics are called for, a short test is always unfair because so much depends upon the chance element in selecting items for a short test.

Let us suppose that the contents of a certain course can be covered in the answering of sixty well-put questions. If only five of these questions be given in the examination, it is not impossible, though it would be rare for such an extreme case to occur, that he who knows only five will get the exact five he knows and he who knows 55 will not get a single question that he knows. There are certain to be a number of students who 'know' only twenty questions who will pass gloriously, while some who 'know' forty will fail. Those who know forty questions are twice as likely to pass as those who know only twenty but one in four lose on such a 'two-out-of-three' choice. It is small comfort to such an unlucky one who did not pass to know that it was luck and not ability that put a numskull ahead of him because the latter was fortunate enough to be asked the few questions he knew.

Every examination has some element of chance in it for the examinee, but in the old type of examination this element of chance is much larger than in the new type which covers much more ground. In the old type of ex-

amination in India the mediocre 50 per cent will probably have their fate decided largely on the chance of getting one certain question rather than another. The change of one question may change many fates.

The author recently collected five primary-certificate examination papers given by various inspecting officers. He gave the arithmetic paper to the fourth class students in five different schools with the object of seeing whether it made much difference just which of these papers a pupil was assigned. There were only five questions in each paper. In the Central Provinces no pupil gets the right to enter middle school if he fails in the primary-certificate examination; so the paper is a very important one. The arithmetic paper is the one which causes most failures. Only 39 per cent of the 125 pupils either passed or failed in all the five papers. This means that only 39 per cent of the pupils were in the class in which it made no important difference which of the five papers was assigned. To the rest it made a tremendous difference. Thirty-seven passed all five examinations but only twenty-one of these thirty-seven succeeded in passing a sixth examination which was later given by the inspecting officer. When comparing the marks each pupil received in all six examinations with the marks he got in one examination alone, it was found that the pupils who had a high average for all the examinations had a high score in any one examination only to the extent of 65 per cent. When any one examination was compared with any other examination, there was an agreement of only 33 per cent as to whom the best pupils were. This shows that one examination of five questions is only about half as reliable as the five examinations taken together. In other words, mere chance decides very much in the primary-certificate

examination, and this could be very considerably reduced if the examination were longer.

It is impossible to make the primary examination as now conducted longer. An hour and a half is given for each paper and every pupil has about four papers. This is plenty for a day's work and the pupils are thoroughly tired at the end of the day. The examining officer has about eighty schools to examine and can devote only one day to each. But is it necessary that such a small amount of material should take so much time? An adult need only sit down and answer a test like *The Sones-Harry High School Achievement Test* to realize that a good part of the education one has received from Class I to the B.A. degree can be tested in less than two hours. No mere windfall of lucky questions can put one through this test; neither is there cause to fear that a competitor in the examination will be much more or less 'lucky' than oneself unless he deserves to be. That makes the test fair and reliable for all.

Third Advantage

They are very easy to score as compared with the older type of examination. Notwithstanding the fact that the tests are more comprehensive, they may be scored in less time and with less labour. This is of itself a sufficient reason to make them attractive to all teachers.

The length of time it takes to score a test depends largely on how well the scoring is organized.

I know of a teacher who complained that scoring True-False tests took a very long time. Watching him score tests, I found that he read every one of the fifty items through on each test paper. That necessarily takes time and is exhausting because of the monotony and eye-strain. The proper way to score a True-False test is to take a

test form and fill it out correctly. Then all that needs to be done is to put one's own filled-out test beside the pupil's paper and note wherein the pupil's copy differs.

All one needs do is to let the eye match the pluses and minuses and mark accordingly. In Standardized tests a scoring key is usually supplied which makes marking easy. When an unprinted informal test is given the teacher should supply lined paper of uniform lining. Then the pluses and minuses on all papers will be the same distance apart, which facilitates matching papers. When it takes a long time to score New-Type tests, a better organization of the scoring is needed. A quick method can be devised for each test separately, depending partly on the arrangement of the items on the paper.

Fourth Advantage

Pupils like them. There is no question as to the accuracy of the marks they receive. There is no chance for the teacher to show favouritism or personal bias. The pupils, relieved of much writing, find these tests less tiring.

One need hardly comment on the fact that children like these tests. Seeing is believing. The advice is 'Try it'. A good New-Type test is an interesting puzzle to boy, girl, man or woman. Some people have the idea that education and examinations must be painful to be useful. The painlessness of a New-Type test must not be construed to mean that the test does not make the examinee work hard and apply himself. It is interest and not easiness that makes the test fun. Easy tests cease to be fun.

A teacher may grade his papers of an Old-Type examination as fairly and conscientiously as is humanly possible. And yet many a pupil will think the examination unfair. Even with safeguards there will be plenty of reason for the pupil to think that his paper was not graded with

the same leniency as Lal Kumar's, whom the teacher seems to like especially. The fact that last time his own paper was graded more leniently than Lal Kumar's does not impress him nearly as much as the fact that, according to his own way of thinking, he got the worst of the deal today. In objective tests this complaint is very largely eliminated. One can safely turn the papers back to the pupils for verification.

Examining boards in India find complaints of unfair or mistaken evaluation so common that examining is done strictly confidentially, the papers containing no names of pupils but only numbers. If students were allowed to see their papers after they have been marked there might be even greater accusation that the evaluation was unfair. As it is, there is complaint and suspicion after the announcement of every year's results, and many a student feels that he has been discriminated against. Objective tests might not relieve all suspicion and complaint, but it is fairly certain that much of it would be relieved and much fertile soil for the generation of suspicion would be removed.

The tedious and tiring part of sitting in an examination is not that it requires thinking and memory work but that it requires putting thoughts into language and writing. Writing is slow work. Almost in a flash, one can think up answers to questions that it may take half-an-hour to write. In the lengthy writing, energy is exhausted that might have been spent in answering many questions. It is true that the pupil must learn to express himself in language and in writing even though it is painful and tiring to do so. The author is keenly aware of this as he sits and writes these few thoughts that come quickly but take such a long time to reproduce in suitable language.

medium, that his thought-trend is continually broken, which is extremely annoying and tiring.

An essay type of examination limits the speed of thinking of the examinee to his speed of writing. One can, if the mind be busy, think infinitely more in half-an-hour than one can write in a day.

There is a proper time for testing expressional work. When we test geography, why not stick to geography and leave essay-writing alone for the time being? Written expression can be tested separately at another time.

Fifth Advantage

They are more educative for the pupils. Let your pupils score their own tests directly after they take them by exchanging papers with their neighbours. It is then that their interest is at white heat. After the papers have been scored, the questions may be discussed. The pupils readily see where they need further study.

Tests may be constructed whose main purpose is that of teaching. Almost any School-subject test can be used for teaching points later on. A short preliminary test is very often an effective introduction to a new unit of work to be begun, especially so if review of former work is needed or if the subject is one in which the material is not entirely new. For instance, suppose that grammar is to be studied. A survey test on the kind of grammar being used brings pupils and teachers squarely into the subject with definite problems before them.

The function of a teaching test will be discussed in a later chapter. The matter is mentioned here only as an illustration of how easily New-Type tests lend themselves to use educationally as well as for inspection and class promotion purposes.

Sixth Advantage

In addition to the five advantages quoted from Symonds we might add several more which apply particularly to the Indian situation.

The first of these now to be mentioned can conceivably be made to apply to the traditional types of tests almost as much as to the New-Type were certain practices reformed, but tradition will probably prevent such reforms from taking place as long as the old-fashioned examinations monopolize the field. We refer to the advantage the New-Type test has in not putting up an arbitrary uniform passing-mark for all examinations.

It is much easier for an examining board to say: 'The passing mark shall be 50 per cent' (or any other percentage mark, for that matter; whether 50 or 33 or 75 is not the point at issue) than to have to wait until all papers have been graded to decide what mark shall be considered 'passing'. It may seem at first thought to be fair to apply the same passing-mark in all examinations but it is extremely unfair as this works out in practice.

It is utterly impossible, without trying out the tests, to construct two examinations that are equally hard. 'Trying out' is exactly what an examining board cannot do on a large enough scale to mean anything with the Old-Type tests, for the selection of its questions is shrouded in absolute secrecy. One year in a public examination in English, 80 per cent of the pupils will pass and in the next year only 50 per cent, or perhaps even less. Recently a Punjab newspaper reported that in a certain town with three middle schools, not a single pupil passed the English paper of the public school-leaving examination, although in all other subjects about the same number passed as had passed the previous year. The passing mark was the same in English in the pre-

vious year as this year, but the examination was far from being the same, we may safely assume. A uniform passing mark is fair only if the examinations from year to year, and in one subject as well as in another, are of uniform difficulty. This they certainly are not. In some subjects the examinations are consistently easier than in others. What school hasn't heard the whispering among students: 'Don't take this subject; everybody fails in it' or 'Take subject No. X, everybody passes in it' The 'everybody' passing or failing is of course an exaggeration, but it is true that some subjects are chosen not because the pupil wants to take them more than the subjects rejected but because the examination is much more difficult in one than in the other. Uniformity in difficulty is not found in the same subject from year to year nor is the average difficulty over a period of years in various subjects uniform.

In a New-Type test the passing mark will not be a certain percentage of the obtainable marks. Whether a high percentage of the obtainable marks or a low percentage is secured may simply show that the test was either very hard or very easy without giving any indication whatever of the ability of one pupil in comparison with others. What does count is the standing of the pupil in comparison with others.

The usual practice in New-Type tests is to list the students according to the number of marks they secure. One can then give the 20 per cent of the students getting the best marks a 'pass' and fail the 20 per cent getting the poorest marks. Those in between will get varying grades according to their standing. Just what per cent one wants to fail is immaterial to the discussion of how to choose the passing-mark. Practice will vary. Whatever percentage one chooses will be arbitrary or can be made

a matter of rule if it is so desired. Somewhere or other in any scheme there will have to be a fence that is arbitrarily thrown up between the 'passes' and 'failures'. But it is fairer to the student when he has been failed because he is one of the 25 per cent who are the poorest students than because he failed to get 50 per cent of the possible marks which only 30 per cent of the students got on the difficult test of this year whereas 65 per cent of the students got 50 per cent marks on the undoubtedly easier test of last year.

As said before, the system described above which is known as the 'percentile' system, is by no means the monopoly of New-Type tests, but can also be used in the old-fashioned examination and is to some extent being used in one or two provinces.

Seventh Advantage

It is less easy for a mere rote-memory student to make a good score on the New-Type test than on the Old-Type. There is a big difference between test and test of both types, and it is quite apparent that the examinations generally given in India today are much better in this respect than those of ten and twenty years ago. Memorization of stock answers to all questions given in the last ten years is not as great a help in the examination as it used to be. But even improved traditional examinations will probably not approach a simple well-made New-Type test in discounting the value of last minute cramming for the examination.

In a country with an examination system like that of India, the questions put in public examinations and by inspecting officers very greatly influence the type of instruction teachers give. In very many cases the actual syllabus of the teacher, no matter what the printed syllabus

bus may be, is what the teacher expects in the examination. The greater, therefore, the need that the examination stress the kind of education that should be given. The reader is recommended to study a fair sampling of Standardized tests in the subjects he teaches or examines to see if these do not stress the type of education that is urgently needed more than the ordinary-type of present-day old-fashioned examination does.

Eighth Advantage

In the case of Standardized tests or any extensively given test, the norms or results furnished enable teachers to see just where one class stands in comparison with others, and how the work of the individual student compares with that of others. This is highly valuable since the experience of an individual or even a single class or single school is not at all typical of what is to be expected generally. Such knowledge is also of value to the student himself, for all of us, if we are to know our strong points and weak points, need to know how we compare with others of our age and opportunities who live beyond the limited little circle in which we move.

Just what norms add to our knowledge of others and ourselves will be explained more fully in the chapters on intelligence testing.

Ninth Advantage

Standardized tests are useful in helping to determine the proper grading of the curriculum. Tradition has assigned most subjects to the classes in which they are now being taught. Some subjects are not well adapted to the age of the pupils we expect to find in the specified class. The most glaring instance is in the case of the teaching of arithmetic. Multiplication and division* with multipliers and divisors of two digits are commonly being taught in

Classes II and III in India. It has been found that the child of average intelligence does not learn these two intricate processes effectively until he is at least ten years old. Teaching these processes in Class II is entirely out of place as can be quickly proved to anyone who will take the trouble to examine three or four thousand school children chosen at random and carefully tabulate the results. The proper class-placement of the subject is quickly revealed as soon as we use the method used in Standardized tests of recording the successes and failures of pupils in each item of the curriculum. This can be done with an ordinary test as well as with a Standardized test, but the point is that Standardized tests almost invariably give us the information as to the appropriateness or inappropriateness of the class-placement of the subject, while this information is rarely sorted out where the Old-Type tests are used. One of the first effects of the extension of modern testing in India will be the demand for a more scientific grading of the curriculum. The information for this scientific grading will become available as soon as we have tests which are properly Standardized.

Standardized tests by no means invariably recommend an easier curriculum. They have in many cases shown that much higher standards can be demanded in some subjects than are at present demanded, and that some subjects are actually too easy for the class in which they are first introduced and should be taught sooner. Scientific treatment of the class-placement of various subjects has on the whole lead to a curriculum that covers much more ground and is broader than the old. School work may seem to be easier to the child than formerly, but the child learns a good many things his forebears did not. Some things are now being taught that children were not thought to be capable of learning at such a tender age in

the 'good old days'. We learn best when we are interested in our work. Interest in school work does not thrive when the work assigned is too difficult for intelligent participation nor when it is so easy as to fail to challenge the pupil to do his best. The Old-Type tests did not help us to find the proper class-placement of the subject matter in the way that Standardized tests do.

The advantages of New-Type tests are several but there are also limitations which are by no means to be ignored. The Old-Type test has some advantages which make it undesirable that it should be entirely scrapped in favour of the New-Type.

Symonds gives two rather complicated accounts of the limitations of New-Type examinations:

First Limitation

The pupil does not have an opportunity to show his ability to organize his thought. He has nothing to do except check over the truth and falsity of the statements or fill in the missing words or incomplete sentences. Pupils miss the valuable experience of making comparisons, giving explanations or giving definitions. They are not asked to summarize the material or to make applications of principles, and of course, these are valuable abilities which we do not want to neglect.

He continues that the 'traditional examination also has strong enough merits to guarantee its continued existence'.

New-Type tests are not meant to supplant entirely the Essay-type. The expression of thought through language and the organization of thought and knowledge is one of the hardest and most necessary things for a serious student to learn to do, and needs to be done not only in the composition hour but in connexion with any subject worth spending time and thought on. It is rightly said that

there is no thinking without language, from which it seems no more than logical to conclude that we have not thought a matter out thoroughly until we have reduced it to an adequate expression through language.

As implied in the previous pages, where the question of fatigue in the writing of long answers was discussed, the point is not of the need of expressional work of the essay-type but of whether the examination time is the correct time to give the pupil this exercise. Some examinations spend 75 per cent of the time on linguistic expression and 25 per cent surveying the pupil's mastery over the subject-matter itself.

Intelligent users of New-Type tests also give the Old-Type from time to time, or still more often, give both types in one and the same examination.

In the case where the examination is not prepared by the teacher but by an examining board or committee, both types could also be included, or either there can be one paper which is given pre-eminently to test those things which the Essay-type test does better than the New-Type. There is no reason why there could not be, in the high school entrance examination, one paper requiring expression through language, organization of material, making comparisons, and the like. This paper could choose subject-matter from various subjects studied during the year and also include discussions on such matters as are not a part of the school course but are of universal interest to wideawake persons.

There should be more thorough-going tests on expression and organization than the essay test of two hours allows. Pupils should be tested on their ability to work out an extended theme or report on a study project which takes not a mere two hours, but weeks, and demands considerable collection of data through interviews, original

work, and library browsing. Such work should be introduced into the course in India much sooner than it is. Rarely does one find it before entering college and all too rarely there. This type of work, on a modest scale, should be done before the end of the middle school course.

The reader may ask what this has to do with the examinations. It has this to do with examinations: it shows, not only that New-Type tests are deficient in expressional activity but that the traditional types are also deficient. Where New-Type tests are used, this deficiency is very often made up by term papers on term projects which are given an importance not less than that attached to the examination, in some cases being accepted in lieu of the examination.

An examining board would of course have considerable trouble in evaluating term papers. Not being acquainted with what the individual class has done, the board would be in no position to evaluate fairly. The judgement of term papers is the distinct province of the local staff.

Second Limitation

These new types of examinations are not diagnostic in that they do not tell where the pupil's reasoning process goes wrong or where he stops reasoning altogether and starts guessing.

A diagnosis is of value only when proper treatment for the ills identified is carried out. A physician who merely diagnoses my trouble as malaria for the sake of idle curiosity and then runs off without telling me what the trouble is and what to do for it is utterly useless. The question of the diagnostic value of a test only arises when there is a 'follow-up' to affect a cure.

An examining board does not diagnose in this way. It merely publishes the results and does not acquaint the individual teacher with his obvious teaching weaknesses

nor the pupil with his defects and how to remove them. It may report to the Department of Education that the mathematics papers generally were rather lacking in such-and-such respect and make a few suggestions for all schools as a whole, but more than such a mass diagnosis is not expected of an examining board.

A teacher who evaluates the papers of his pupils but does not use them to analyse the individual pupil's problems and carry out, with the help of the pupil, the removal of the defects revealed, is not taking advantage of the diagnostic features that any test may have, whether Old-Type or New-Type. Unfortunately many teachers confine the examinations they give solely to 'passing' or to 'failing' pupils or to trying to find out what the prospects of passing in some coming examination may be. Thus full advantage is not taken of the superior diagnostic values of Old-Type tests.

Those working with New-Type tests have recently paid considerable attention to increasing their diagnostic value. In some subjects special tests are obtainable that have diagnostic value far beyond that of Old-Type tests. In other subjects this is not quite possible. More information on diagnostic tests will be given in a later chapter. Where individualized instruction is not taking place, the diagnostic value of all types of tests is almost entirely ignored.

In addition to the above limitations some common objections to the use of New-Type tests may be discussed. In India the objection is very commonly raised against Standardized tests that once a test had become known it would be utterly useless, since teachers would coach their pupils on the test, thus invalidating it for further use.

If there were only two or three standardized geography tests in a given language, an examining board would

certainly not want to use them because of the unfair advantage secured to pupils in schools where the teacher has coached for the test. If there were ten usable tests in existence, the comprehensiveness of a Standardized test is such that it would be far easier for a teacher to teach his subject thoroughly than for him to coach pupils to steer their way through such a labyrinth of test questions. If under such circumstances a geography teacher digested the contents of all the Standardized tests and used the result as a kind of syllabus in his teaching, the outcome would certainly be a gain in balance and proportion and comprehensiveness over current practice. But when used to prime for an examination the danger to the pupil of getting mixed up hopelessly would far outweigh any likely advantage to be gained from coaching.

An examining board would probably not want to use an already Standardized test. Where thousands are examined, as in the public school-leaving examinations, a new battery could easily be made up each year without much more trouble than that involved in making up the questions under the present system. These would not have to vary so very much from year to year to make intensive coaching on previous years' tests more of a trap than a help.

Standardized tests are generally used for purposes quite different from those of the public examination. They are used for the valuable information supervisor and teacher get from them regarding the state of teaching and about the individual pupil's ability. A teacher might want to cheat the supervisor and make him think the class is better than it actually is, but this he attempts to do with the Old-Type tests certainly as much as, if not more than, with the New-Type, and with better prospects of success. We assume that supervisor and teacher co-operate. A

teacher may succeed in deceiving once or twice but stands little prospect of continued success in deception when New-Type tests are used.

In America and England most of the giving of Standardized Achievement tests is either done by the teacher himself or is a co-operative affair with the examiner. Under such conditions a teacher could only cheat himself. Public school-leaving examinations in India are conducted in an atmosphere of suspicion. The teacher is considered to be either too untrustworthy or incapable of examining his own class. In Standardized tests the original purpose of the test is so entirely different from that of the public school-leaving examination that the same state of tension between teacher and examiner does not exist. It is assumed that the teacher is sincere in his desire to find out the true state of affairs in his class.

The author has used Standardized arithmetic and reading tests in the primary school to determine promotions in the lower three classes. Although he had a choice of only two reading tests, he found little difference in results on these tests between the years after the teachers were acquainted with the tests and the first year of their use. Some teachers had undoubtedly used them for coaching purposes but little children cannot be relied upon to answer sixty items correctly from memory if they do not know how to read well enough to guide them through the test. If they can read well enough they can pass the test without coaching. As regards the arithmetic test, since almost everything the pupil has been taught in computation is asked for, the feat of memorizing the answers demands as much as learning to do the computations. There is no gain in coaching. Teachers decide that it is just as easy to teach reading and arithmetic as to coach for such examinations.

The best safeguard against cheating on the part of the teacher is the development of the scientific spirit in him. This spirit is fostered by scientific testing. Any teacher who is interested enough in his work to make a good teacher becomes as much interested in a fair objective standard as the examiner. With a share of the responsibility of examining he shares the desire for learning the truth about his class.

It is often claimed that the widespread use of Standardized tests in the various school subjects is financially prohibitive. The writer has found that Standardized tests of eight pages can be printed in quantities at one pice per test. Thus an anna examination fee would provide Standardized test papers for four or six subjects, depending on their length. Probably school or pupil are paying as much as this for cyclostyled examinations of the ordinary type and for paper or blank copybooks. Standardized tests can be provided practically as cheaply as the present type of examination if composed and published by non-profit-seeking examiners.

In conclusion

This extended discussion of the relative merits of New-Type and Old-Type tests is of necessity somewhat 'one-sided' as we are comparing tests improved as we would like to see them, with tests as they actually existed before widespread dissatisfaction with them inspired considerable experiment. In other words it is a claim that the attempts to improve examinations has resulted in some real improvements.

This improvement, where it is a real improvement, does not come of itself simply with the use of novel forms described in Chapter I. A New-Type test can be constructed without using a single one of the more novel

forms and a test using the novel forms may be almost entirely devoid of the advantages claimed for the New-Type tests. Making a good test along either the newer or older models is an art and unless certain principles of construction are observed the test will be a botch whether New-Type or Old-Type.

There is no use making New-Type tests just in order to do the 'modern thing'. Actually we don't care what type a test is provided it is a good test. A good test, whether Old-Type or New-Type, will observe among others the following principles.

1. It will be general and comprehensive enough to test thoroughly the pupils' mastery of the desired skills or grasp of a subject matter. It will make it impossible for a pupil to get a good mark just because he got a few 'lucky questions' and will not fail a good pupil just because of a few 'unlucky questions'. Very few examinations in India are adequate in this respect.

2. It will reliably grade the pupils into at least six to ten (or more) different classifications according to their ability.

3. It will be as objective as practicable. Some questions of a discussional nature can be given to advantage but there should also be some work which pins pupil and examiner down to answers which can be quite objectively scored. Even in the essay type of examination the objectivity can be greatly increased with study and care.

4. Other things being equal, the test which examines most intensively or extensively in the least time and with the least fatigue on the part of the pupil is the best.

5. The test should encourage creative and self-reliant work and discourage mere mechanical rote-memory work.

6. The test should encourage the pupil to put forth his best effort.

7 The test should convincingly reveal to the pupil his deficiencies and encourage him to remove them.

8 The test should examine exactly what it pretends to examine and not handicap pupils for deficiencies in unrelated subjects. The more a test tries to examine in subjects extraneous to the avowed subject the less efficiently it examines its own.

9 The standards of a test should be based on actual performance and not upon mere opinion of what a standard should be or on an arbitrary percentage mark.

10 The more a test makes it possible to compare the performance of a certain group of pupils to that of other pupils the more revealing it is.

11. Keep in mind that no test is infallible. The daily work of the pupil and an observation of his habits of work are equally important in appraising the work and ability of a pupil

12. Keep in mind that poor results on the part of a group of pupils indicates a faulty test or one that is not graded to the ability of the pupils or deals with matters the children have not been taught.

V

THE USE OF DIAGNOSTIC AND INSTRUCTIONAL TESTS

REFERENCE to diagnostic tests was made in the previous chapter. This type of test, though decidedly a schoolroom test, differs radically in purpose from the tests hitherto described. The purpose of a diagnostic test is not to classify the pupil according to ability, effort, or training, but to inquire into the nature of the difficulties which a weak pupil is having with his school subjects so that proper remedial instruction may be applied. This purpose will be described in fuller detail in the succeeding pages.

The writer once gave a problem in multiplication to a class and half the pupils attempted a solution as follows:

$$\begin{array}{r} 325 \\ \times 25 \\ \hline 1625 \\ 650 \\ \hline 2255 \end{array}$$

It is quite plain what the trouble of these pupils was. Their trouble was very specific. What they needed to be taught was not arithmetic in general, not even the multiplication process, but the principle involved in multiplying by units, tens, and hundreds. Half an hour's work today and five minutes attention for a week on this specific principle would do more for the pupils than reviewing and re-teaching the entire process for many days. The teacher of this class had seen this same little mistake made day after day for several months but his only

remedy was 'more work in arithmetic'. He gave problem after problem to the pupils to do but did not get down to the specific ailment that needed special treatment. After all, a splinter under my finger nail needs a certain minor operation and not a general physical overhauling; and so a specific ailment in a mental process needs a certain minor operation.

This illustration is an extreme case in which an untrained and, in this case, unintelligent primary school teacher is the actor. But unfortunately, in a less crude way, an almost as flagrant lack of attention to specific difficulties of the pupil is sometimes found in some of the better middle schools. Any teacher who does not analyse in a systematic way the weaknesses of his pupils is quite likely to assume that some of his pupils are generally weak when a diagnosis of their troubles would reveal only a few minor difficulties.

Teachers often spend much time in teaching over and over again what is already known and neglect to teach a few essential details. The neglect makes it appear as if the pupils are generally weak. A certain teacher spent three months teaching 'interest' problems with disappointing results. Closer analysis of the work of the pupils showed that they knew the method of computing interest but were so weak in division and multiplication as to impair greatly their accuracy. This teacher erroneously presumed that they did not know the method of finding the amount of interest. The method had actually been taught much longer than was necessary. The real trouble was of an entirely different order.

Pupils are often accused of being weak in geography and history when their real trouble is weakness in reading. They read so poorly that they cannot study properly and hence do not get the geographical or historical informa-

tion that they could get without difficulty were they able to read adequately.

The means adopted by teachers to remedy common faults are often most devious, as witness the attempt to teach spelling through dictation. Even in the middle school one finds teachers in India dictating long passages from books for the pupils to write as language exercises. Just what might one expect the pupils to learn from this? To write correctly what the teacher reads requires ability to spell on the part of the pupil and some knowledge of phonetics. But why should a teacher dictate an entire paragraph containing 95 per cent words that he knows the pupils can spell in order to test the few words about which there is doubt? An analysis of the pupils' written work will soon reveal which words they most need to learn to spell. These are the words to practice and not entire paragraphs. If the pupils do not use certain words that the teacher wants them to use in written exercises, he can dictate these certain words one by one, or can give written work in which these words are bound to appear. He can then see that they are used again and again. By dictating certain paragraphs today and certain paragraphs tomorrow there is no 'follow-up' in the use of the same words with the result that the practice in spelling is so diffused as to be almost useless. It is best to concentrate on words that need drill and to give enough specific drill to make the correction as permanent as possible.

It is doubtful whether spelling is taught by dictation exercises. Most Indian languages are so phonetic that he who knows his phonetics and has a good ear should invariably write the word correctly provided it is pronounced correctly. If the hearing is not good or knowledge of phonetics is weak, the dictation of entire paragraphs is the poorest method of correcting this. Specific phonetic

drill is needed and practice in the writing of the correct symbols for certain sounds. It will probably be found that only a few of the sounds are confused, and these are the only ones that need drill-work. Where the ear simply does not distinguish between certain sounds there may be a physical handicap which needs medical attention or is irremediable. What concerns us at present is the fact that there is much teaching that scatters its efforts over a wide field without sharply defined aims. Such a method causes much waste of teaching and learning effort. The habit of analysing the weaknesses of pupils in order to determine the specific remedial work needed will put an end to aimless teaching.

No physician considers that he is making headway in a cure until he knows the patient's specific ailment. If the patient recovers without the physician having discovered what the ailment is, he knows that the patient probably grew better by himself without the treatment having expedited matters. Any pupil who recovers from a fault without the teacher having discovered what the nature of the fault was, was cured in spite of the teacher and not because of any help he got unless the help was an accidental 'stray-shot' suggestion. The physician knows that identification of the trouble is more than half the battle won in very many cases. In the teaching profession it is no less important to know exactly what the pupil's trouble is, for the help that can then be given is usually many times more effective than it is when the pupil is simply considered weak. The time will come when a teacher will be as much ashamed to answer 'weak' to an inquiry about the condition of the pupil as a doctor is to write in a medical report 'sick'.

Any teacher is well advised to begin the year's work with a new class by giving a few short tests to find out

just what the strengths and weaknesses of his pupils are. No teacher can make a success of his job if he takes the attitude: 'The pupil is weak in reading but that is a primary school subject. In the middle school we are not responsible for that subject.' Equally out of place is the reaction 'Multiplication and division are Class III subjects. We have no time for them in the middle school.' If the foundation is weak it needs to be repaired regardless of who built it. A good teacher teaches pupils first and the curriculum second. That which the pupil specifically needs in order that he may be in a position to understand today's lesson is of greater importance than teaching the lesson appointed for today itself. No teacher is prepared to start the year's work with the pupils until the foundational learnings necessary for the course have been examined and repaired where necessary. All teachers are well advised to examine their pupils in that knowledge which are presupposed for the new course.

Assuming a teacher is beginning mathematics in Class V, he should first give a comprehensive survey test including computation in the four simple rules and the four compound rules, simple fractions, the use of the decimal, and whatever other processes the pupil is supposed to know. His test should include several examples of each process that the pupil is expected to know. The teacher should then analyse the test paper and write down the name of every pupil who did poorly in addition, in multiplication, or in any other process together with the process the pupil is weak in. He will then have a good idea of who the weak members of the class are and what the weaknesses of the class as a whole are. He may have to review a few processes for the entire class. In some of the processes there will be five to ten pupils who need special work. Inform each weak pupil what his weaknesses are,

and let him do special home-work. Ten minutes of each class period during the first two or three months used to supervise the special assignments for weak pupils will save the middle-school career of many a pupil. As it is today, many middle-school teachers start right in teaching the new course when the foundation on which the new course intends to build is too weak to support further structure, but could be put into a proper state of repair with a little skilful personal attention on the part of the teacher.

When a certain pupil is weak in multiplication, need the entire process of multiplication be retaught? Certainly not. Find out where the pupil's difficulty is. This can be done by giving him the following test:—

1. 21×5	51×8	82×3
2. 22×5	52×8	82×6
3. 462×7	189×9	189×4
4. 380×5	405×6	306×7
5. 24×11	33×22	41×84
6. 48×48	63×31	98×57
7. 513×92	789×18	316×54
8. 595×60	505×23	3007×40
9. 473×708	395×216	9045×350

Any pupil who can do up to row five in the above test satisfactorily but does not do well with row six and utterly fails from row seven onward evidently needs help with examples of the difficulty of row six onwards. Each row denotes progress of one teaching step. We now know at what step to begin repair work. Let the teacher make the pupil work examples of this difficulty on the blackboard and instead of letting him do it silently let him do his reasoning aloud. Hearing the pupil do his reasoning will show the teacher the faulty reasoning processes like an X-ray shows up certain internal defects to a physician.

When the faulty reasoning has been revealed, the application of the remedy is usually not difficult. The thing is to locate the faulty reasoning. A large amount of explaining and re-teaching will be practically useless if we do not locate this.

Analysis of pupils' work has revealed the following errors to be quite common in multiplication¹ :—

MULTIPLICATION: (*Place a check before each habit observed in the pupils work*).

- | | |
|---|--|
| — Errors in combinations | — Error in single zero combinations, zero in multiplicand |
| — Error in adding the carried number | — Confused products when multiplier had two or more digits |
| — Wrote rows of zeros | — Repeated part of table |
| — Carried a wrong number | — Multiplied by adding |
| — Errors in addition | — Did not multiply a digit in multiplicand |
| — Forgot to carry | — Based unknown combination on another |
| — Used multiplicand as multiplier | — Errors in reading |
| — Error in single zero combinations, zero as multiplier | — Omitted digit in product |
| — Used wrong process— added | — Errors in writing product |
| — Counted to carry | — Errors in carrying into zero |
| — Omitted digit in multiplier | — Illegible figures |
| — Wrote carried number | — Forgot to add partial products |

¹ *Diagnostic Chart for Individual Difficulties in the Fundamental Processes in Arithmetic* by G. T. Buswell and Lenore John. Public School Publishing Company, 1925.

- | | |
|--|----------------------------------|
| - Omitted digit in multiplicand | - Split multiplier |
| - Errors due to zero in multiplicand | - Wrote wrong digit of product |
| - Error in position of partial products | - Multiplied by same digit twice |
| - Counted to get multiplication combinations | - Reversed digits in product |
| | - Wrote tables |

Knowing which of these errors the pupil makes, gives the teacher definite suggestions for remedial work. Of course it may be that a pupil makes several of these errors quite commonly. In that case more or less the whole process must be retaught; but often only one or two of these errors are commonly made and this fault can be removed with a little individual attention on the part of the teacher.

There are similar tests to be had for addition, subtraction, division, fractions, decimals, etc. The well-known *Compass Diagnostic Tests*¹ take one through twenty processes step by step, helping one to find the level of efficiency of the pupil in the various processes. All of these tests would not be given to one person, but only the test in those processes in which the pupil is known to be, or is suspected of being weak.

The above error schedule given for multiplication is so useful that a teacher of mathematics is well advised to purchase a copy of the Buswell-John tests in order to have the error schedules before him not only in multiplication but the other fundamental processes.

In problem solving, the difficulties are of quite a different nature than in computation. *The Stevenson Arithmetic*

¹ Edited by G. M. Ruch, and published by C. A. Gregory Co.

*Reading (Problem Analysis) Tests*¹ determine the ability of the pupils along the following lines:

1. To read the problem.
2. To analyse the problem and determine what facts are given and what is to be found by the solution.
3. To judge approximately what answer is reasonable.
4. To determine the process to be used in solving the problem to get the exact answer.

An example from the test is given herewith²:—

On May 5th, Alice deposited \$ 0.50 in the school bank, on the 10th she deposited \$ 1.50, on the 15th she put in \$ 0.50 and on the 20th she deposited \$ 1.00. How much did she deposit altogether during May?

A. Which of the following facts are given in the problem?

1. The different amounts deposited.
2. The total amount deposited.
3. The interest paid by the bank.
4. The time when money was withdrawn.

B. Which of the following things are you asked to find out in the problem?

1. The profit gained on the deposits.
2. The number of times she deposited money.
3. The total amount deposited.
4. The amount of each deposit.

C. Which is the most reasonable answer?

\$22.00 \$15.50 \$1.00 \$3.50

D. Which process should be used in solving the problem?

Addition Subtraction Multiplication Division

¹ Public School Publishing Company

² Printed here by courtesy of the author, Dr P. R. Stevenson, and the publishers, the Public School Publishing Company, Bloomington, Illinois, U.S.A.

F. J. Schonell in *The Testing of Intelligence*¹ gives a suggestive schedule of reasons for failure in solving problems.

Reasons for Failure in Problem Solving

(i) Computational inaccuracy. (ii) Inability to read the problem intelligently. (iii) Inability to select relationship between fundamentals in the information: (a) Inability to select what is given from what is required, (b) Inability to select pertinent facts for working. (iv) Inability to select the proper process. (v) Carelessness in reading the problem (misreading of statements or numbers). (vi) Premature generalization due to catch-words or signs (vii) Failure to try the problem due to large numbers or its detachment from child's knowledge and interests. (viii) Failure to make a simple mental example and find an approximate answer.

Arithmetic is one of the easiest subjects in which to do diagnostic work. Reading is much more difficult. There are some few types of reading difficulties which cannot be properly disclosed without optical instruments and the skill of an expert. In reading, the eye does a very intricate piece of work, and a thorough knowledge of the physiological process taking place in reading is required for diagnosis of faulty eye-work.²

Fortunately only a part of the reading difficulties commonly met require such expert care. There are many common difficulties which the ordinary teacher who informs himself on the subject can be helpful in removing. Just what can be done is indicated by the description of two tests by the publishers of the following tests.

¹ Edited by H. R. Hamley and published by Evans Brothers, London, 1935. The chapters by Schonell on Diagnostic Tests are well worth a teacher's careful study.

² See chapters II and XVI of *Suggestions for the Improvement of Reading in Indian Secondary Schools*. (In the press)

*Pressey Diagnostic Tests in Fundamental Reading Habits*¹

These are individual tests for measuring eye movements, vocalization, and vocabulary. The teacher observes the number of fixations, regressions, etc., and makes a record of them while the child is reading.

Research has shown that poor eye movements, extreme vocalization, and inadequate vocabulary cause children to read poorly and that deficiencies in reading are due to one or more particular 'bad habits' which may be located and remedied. It has also been shown that progress in any subject comes most quickly when specific errors in the fundamental habits are located by analysis and remedied by instruction directed especially against the particular errors found.

Doctor Pressey has brought together the outstanding results as regards fundamental reading habits in such a way that the average teacher can make a rough analysis of difficulties with very simple equipment. The only materials needed for this work are: (1) A small mirror, (2) the appropriate folder of reading sections, (3) a book against which to hold the mirror and folder, and (4) the pupil's record card. The analysis includes three simple measurements of eye movements, one of vocalization, and one of certain elements in the acquisition of vocabulary.

*Pressey Diagnostic Reading Tests for Classes III to IX*²

The words of the Vocabulary test were selected from the Thorndike list of the most common 10,000 words, and are so arranged that by simply adding two zeros to a pupil's score, a direct statement is obtained of the size of his reading vocabulary. Thus, if a child scores 30, this indicates that he knows about 3,000 of the 10,000 commonly used words.

The paragraph Meaning test requires the pupil to mark topic sentences to a variety of carefully organized paragraphs. It requires consideration of each paragraph

¹ By S. L. and L. C. Pressey and published by the Public School Publishing Company. ² By L. C. Pressey.

as a whole; gets away from artificiality and emphasis on unimportant details; and permits sixty paragraphs to be used without undue lengthening of the test, with resulting better reliability and differential efficiency.

The Speed test is composed of a simple, 3,300 word narrative, in which very easy and quickly-marked questions are embedded. Since the questions are inserted in the reading material, it is hardly possible for a child to look for questions without reading the material through consecutively. The pupil's score is made up only from material which he actually did read—and that with some understanding.

The publisher's claim for the two above-mentioned tests does not seem to be exaggerated. There are diagnostic tests by other publishers which are probably equally good.

An ordinary well-constructed test, standardized or informal, can also be used to some extent. With the Hindi reading test for primary schools described in Chapter III the teacher can quickly tell whether the main trouble of the pupil is in paragraph meaning or whether it starts lower down in the scale with word recognition, or in reading simple isolated sentences like those in the part on following instructions. As the sentences and paragraphs are graded in difficulty one can, by scrutiny of the paper, quickly see on what level of efficiency the weakness makes itself strongly felt. The test for middle schools can be similarly used. The use of tests in this way hardly merits the dignity of the name diagnostic work, but it is at least preliminary to it and perhaps the most practical procedure for the ordinary teacher.

At this point it might be well to explain the difference and similarity between Instructional and Diagnostic tests. An Instructional test is of the type a teacher would give when he has finished a unit of work and wants to see whether pupils have mastered it before going on to the

next unit. He gives such a test because he wants to make up the deficiencies he finds. A Diagnostic test pays little attention to what has just been taught but reviews the whole subject to catch up any difficulty that the pupil may have in any phase of the work. It seeks to discover the cause of the trouble. It is more comprehensive and thorough than the Instructional test. What both of these tests have in common is that they seek to find out what weaknesses to correct. It is this similarity that concerns us more than the difference between the two. Both of them aim at giving information needed for remedying weaknesses. In this chapter we have rather loosely used the name Diagnostic test for Instructional test also, as many distinctions are apt to confuse the novice in New-Type testing. The one point the author wishes to emphasize is that both are tests that can be used profitably to remedy instructional defects and individual weaknesses of pupils.

One can get diagnostic and instructional tests in all the principal school subjects, including writing and language usage. In addition one can get very comprehensive analyses of the common faults of pupils with careful case studies including an account of the effects of various types of remedial instruction. Some of the best research work being done in education at present is being done in these phases of instruction. Education is seeing the futility of teaching only the class as a whole and is emphasizing at least some individual instruction for those in need of it. The classroom is not different from life in general where all people simply cannot be herded together and treated like one composite individual. There are vast differences between person and person both as regards capacity and quality.

Good teaching simply cannot ignore the many problems of the individual student. Ways of teaching are being

developed which make it quite possible for the teacher to teach both the class as a whole and the individual. It cannot be said that there is no time for individual instruction, for to ignore the difficulties of the individual wastes time for the class as a whole. It is a matter of proper organization of the time available. A method of part-time class instruction and part-time supervision of the pupils as they do their own work is finding increasing favour. Properly prepared teaching materials, such as work-books, make this easily possible.

It is not necessary to have a printed diagnostic test in order to do diagnostic work. Most diagnostic work is done with short informal tests that the teacher has made or adapted. Merely sending a pupil to the blackboard and making him work his computations aloud is diagnostic work if the teacher is alert to analyse the pupil's method. What makes a test diagnostic or otherwise is the way the teacher uses it. Of course some tests are much better than others for this purpose but no test is diagnostic simply because the title contains the word 'diagnostic'. Only when it is used to analyse the difficulties of the individual does it become truly diagnostic. One can diagnose the troubles of the class as a whole but there is a limit to what can be done in this way. The doctor can treat a family as a whole only occasionally. He is soon forced to put his stethoscope on the individual. So also is the teacher.

A teacher who has obtained for suggestion and guidance a few printed diagnostic tests in his special subjects should be in a position to make up informal exercises which are helpful. When a sample of tests in a certain subject is ordered from American publishers, one or two diagnostic tests will certainly be included.

Diagnostic examination without remedial instruction to

follow is as useful as an examination of the kidneys which reveals pus cells is, if nothing is done to remedy the obviously bad situation. There is no value in diagnostic tests if the teacher does not intend to work with small groups or individuals to remedy their defects.

What treatment is necessary when defects are revealed? This depends entirely upon the defect. Sometimes the remedy suggests itself as soon as we have located the defect. At other times the remedy appears to be a task almost as formidable as finding the cure for cancer. A good diagnostic test will contain suggestions to teachers for remedial work but endless ingenuity and study of his cases is nevertheless required on the part of the teacher. But that is exactly what produces good teaching.

In India attention to the individual has suffered tremendously because of the large amount of elimination of pupils in each class. A pupil who is weak in arithmetic is not long tolerated in the school. He is failed so often that he discontinues school. A pupil weak in reading is failed or learns to memorize his material. In America some schools employ teachers specially trained to coach pupils weak in reading so that they can compete with their classmates. It sometimes happens that quite intelligent boys and girls have not learned to read. In America this is considered as bad a handicap in life as hernia; treatment is considered just as necessary as in the case of the physical defect. In India, as the number of illiterates grows less and the handicap of a person who cannot read because of some slight remedial defect comes to be considered a needless tragedy, more attention will be given to remedial work. In other subjects we may expect similar treatment. The day may still be far off but must surely come when we will not permit nearly half of India's childhood, which finds its way to the school, to be weeded out in Class I,

and a high percentage in every succeeding class. It is not necessary to fail every child who reveals a weakness in his scholastic learning. Many situations of academic ill-health may be remedied, so why throw them out to die academically, by failing them, until they are swept out of school by the force of circumstances which could easily have been remedied at the proper time?

Diagnostic tests may be constructed for all levels of development, from very simple informal tests to elaborate standardized ones which inquire into the standing of the individual in comparison to others of the same age and school experience. Diagnostic features may be included in School-subject tests, such as are described in Chapter III, thus making the test usable in two ways. As said before, it is the use to which a test is put which determines its diagnostic value. This does not mean that diagnostic features are more or less accidental features of any test. If claim is made for a test that it has diagnostic features it should have an extraordinary quality that lends itself to utilization in this way. The author of the test must have successfully directed much care and attention to the development of this quality in the test.

PART II

PSYCHOLOGICAL TESTS

VI

WHAT ARE INTELLIGENCE TESTS?

WITH intelligence tests we come to those types of tests which may be termed psychological because to a much greater extent than school tests do, they concern themselves with the way the human mind responds. School tests are less concerned with the way the mind responds than with checking up to see whether or not a pupil is learning what he is supposed to learn and how well he is doing it. Psychology has no set programme for the pupil like the school has and so its interest is not in the curriculum but in the learning process or in any kind of nervous reaction. Psychological tests may be expected to conform to psychology's interest in the working of the mind and its inseparable partner the body, rather than in diplomas and the mastery of subject-matter.

Psychological tests attempt to do a number of things which, if successfully carried out, are of no little importance to the educator. They claim to be able to measure the mental ability of human beings (Intelligence tests). They are trying to measure mechanical ability and other abilities some of which are very much dependent on general intelligence and some of which are not (such as Mechanical Aptitude tests). They are trying to perfect a technique by which they can predict success in various kinds of work, yet to be learned, on the basis of the pupils' performance at present (Prognostic tests). Then there are the attempts to measure attitudes, emotions, temperaments, etc. These will all be considered briefly.

Individual items in the various kinds of tests may resemble each other greatly or even be identical; but the purpose behind asking these questions is quite different in each type of test. A good examiner sets his questions as shrewdly as a good cross-examiner. Both prosecutor and defence lawyer in a trial may ask similar questions but they intend to use the answers in quite different ways. In different types of tests similar questions are asked to give information on quite different types of mental reaction. It behoves one to know exactly what a certain test is trying to do and how it expects to do it before one can use the results of such a test. In the discussion of the types of tests to follow, the emphasis will be less on the content of the test than on the purpose, for that is of prime importance.

Our question for this chapter will be: What is the purpose of the Intelligence test?

Every teacher has noticed that although all of the pupils in his class have approximately the same opportunities, some of them learn much more quickly than others. Sometimes a certain pupil with less opportunities than the others is, in spite of his handicap, the fastest learner, while one with the best opportunities lags near the bottom of his class. Interest and effort in school work may have something to do with it but not infrequently the hardest workers get along least well and a certain pupil who practically 'never works at all' and does not care whether he passes or not stands at the top of his class. What is it that enables one pupil to learn more easily than another? This quality is called 'intelligence'. One was born with much more potential ability than the other, which shows itself in the greater ease with which he can learn.

A pupil's success in school depends on a number of

conditions, of which the following are the most important:

1. His native intelligence.
2. The length of time he has attended school.
3. His health.
4. His industry.
5. The quality of the teaching he receives.

We can easily find out how long the pupil has attended school; a physician can usually tell us what his state of health is, the teacher usually knows whether he tries hard and it is not difficult to find out how good the teaching has been. The most difficult thing to find out is what native intelligence he has. There is little in the pupil's life that either helps him or hinders him more than just this factor; the amount of intelligence he inherits. No matter what may be done for a child and no matter how hard he works, a feeble-minded child can never be a brilliant adult. On the other hand a bright child can be confronted with many handicaps such as would seriously hold back the average child and still come out far ahead of his companions. It often happens, however, that a brilliant child is not recognized as such until he has grown up and distinguished himself.

A good home and environment, opportunities for education, and hard work, are all tremendous aids to success. But if we happen to have been born abnormally slow learners, all that teachers and elders can possibly do is to improve the situation only slightly. If, on the other hand, we are amongst those whose learning ability is above normal, and we receive and make use of a fair opportunity, we are destined for distinction. Even before we are born it has already been settled, other conditions being about normal, whether we are to be among the distinguished, mediocre or unsuccessful. Already it has

been settled whether there is human possibility of pushing us up to B.A. standard or whether coolie rank is the best that can be expected from our efforts no matter what educational opportunities are made available. Extreme cases have been cited. Fortunately the majority of humans are so near normal that hard work is the determining factor in our success.

Sixty per cent of us (judging by data from England and America; there are no data for India) are so near to normal that what is expected of the average man may be expected from us. Another fifteen per cent are gifted so far in advance of the average that mere normal people have to work much harder in order to accomplish what these can do with relatively little effort. Another fifteen per cent are so far below normal that they are definitely out of the race in anything demanding high grade intelligence. Or, since there are degrees of intelligence from that of the lowest rank of idiot to that of the rarest genius, we might say, about fifty per cent of the people can be expected to do better than average work and fifty per cent less than average work, by far the greater number of us just slightly better or poorer than average. Where we belong in this line-up has been determined roughly, barring accident or peculiarly unfortunate environment or wasted opportunity, at the time of our conception, although our later success is by no means wholly dependent on our position in this line-up.

It would be extremely helpful if it could early be found out just what the possibilities of a child's later development are. It is wasted effort to try to push the subnormal through college and it is wasted opportunity for society to let its most gifted people spend their time and talents doing what the ordinary coolie can accomplish. Intelligence tests attempt to find out what the native endow-

ments of the individual are. We do not say that fate decides what we shall amount to regardless of what effort we put forth. Hard work, stable character, and worthy ambitions and aims accomplish wonders, while brilliance without these qualities often ends in tragic failure. But, if we can find out even approximately what native ability a child has, we may be able to keep him from wasting his time on that which is beyond his capacity while neglecting the opportunities which lie within his capacities.

Measuring intelligence is much harder than measuring achievement in the school. When we measure achievement in school we do not try to separate what the pupil learns by himself because he is an intelligent lad, from what he has learned because of the teacher's infinite patience in drilling something into him. We simply try to find out what he can do, no matter who or what is responsible for his having learned to do it. In testing intelligence, however, we have to isolate to the utmost possible extent native ability from other influences and seek to measure that alone.

The man who launched intelligence testing on the career which won for it recognition from the scientific world, was a Frenchman, Alfred Binet. This man was a real pioneer in science. Binet was connected with a school for mental defectives. He pondered over the question as to what his mental defectives did not have that normal children do have. He noticed that mental defectives are different from normal children in that they have in smaller degree what normal children have in a larger degree. He noticed that some mental defectives could be taught much more than others, which suggested to him that it would indeed be helpful if he could construct some simple device that would tell him how defective or how nearly normal his various

charges were. From this he expected to know approximately what things might be done for them and what things were probably a waste of time. Since the mentally retarded learn so much more slowly than others, much time was usually wasted before it was discovered what is beyond the pupil's capacity and what is within his capacity. His desire to perfect a device to give him better information on these points led Binet to experimentation which resulted in perfecting the first acceptable intelligence test.

Binet also discovered that a teacher of normal children has a much easier time in adapting his teaching method to the age of his pupils; for the teaching profession has gradually built up a tradition of how to deal with children of the various ages. The teaching profession has made many glaring mistakes in its estimate of the ability of children but on the whole the estimate has been a fair approximation. But in the case of mental defectives the age of the child gave little intimation of what was to be expected of him. A ten-year-old frequently has to be taught like a five-year-old, or like a seven-year-old, or even like a three-year-old. But which? Binet longed for some way to find out early in each pupil's career just how each pupil should be treated. It was by considering the mental defective to be mentally like a much younger child that gave Binet his idea of measuring intelligence by comparing the activities and abilities of the children of one age with those of other ages.

Binet received suggestions from the work of Galton, Wundt, Cattell and others who had measured the speed and accuracy of seeing and hearing, strength of memory, ability to recognize various shapes and forms, strength of hand grip, attention span, and so forth. Without the suggestion of these experiments it is doubtful how much

Binet would have accomplished. But Binet early felt that intelligence is much more than just a matter of sight, hearing, muscle control, memory and concentration. He wanted to know how one could measure the 'richness of inspiration, accuracy of judgement, and general ability of the mind'. He did not try to break up intelligence into many parts to be measured separately. He sought to measure the combined effects of attention, imagination, judgement and reasoning as the human being shows these qualities in doing fairly complicated tasks. For many years he experimented with many types of mental and physical responses that showed not the knowledge of the pupil but his ability to do novel tasks revealing imagination, judgement, etc.

During all these experiments Binet again and again asked himself: 'What is the normal child of various ages able to do?' He felt that if he could learn what the normal child could do, he could classify the mental defectives by comparing them with normal children.

But who is the normal child?

How can we tell what the normal child can do if we do not know who the normal child is? Which of the hundreds of children we know is normal? Psychology had already progressed so far as to suggest to Binet that, by testing hundreds of children of a certain age, taking care that these were chosen at random and were not a select group, the average performance of the entire group would be a truer indication of what the strictly normal or average child is than that provided by any one child. Binet, together with his friend Thomas Simon, worked on this idea for many years until in 1908 they were able to publish a list of the things one might expect the normal child of various ages to do. Here is the list which rightly earned for Binet the title of 'Father of Intelligence Testing', and at

last provided a crude yard-stick with which to measure intelligence according to a definite standard.

AGE 3

1. Pointing to nose, eyes and mouth.
2. Repetition of short sentences
3. Repetition of two digits
4. Enumeration of objects in pictures
5. Knows his last name

AGE 4

1. Knows sex.
2. Names familiar objects.
3. Repetition of three digits.
4. Knows longer of two lines.

AGE 5

1. Compares two boxes of different weight.
2. Copies square.
3. Game of patience Rectangular card cut diagonally to be reconstructed according to a similar uncut card placed before the child.
4. Counts four coins
5. Repeats a sentence of ten syllables.

AGE 6

1. Knows right and left.
2. Repetition of a sentence of sixteen syllables.
3. Aesthetic comparison Chose the prettiest of three pairs of faces.
4. Definition of familiar objects
5. Executes three commissions
6. Knows age.
7. Distinction between morning and afternoon.

AGE 7

1. Unfinished pictures. Tells what is missing.
2. Number of fingers. Knows number on each hand and on both hands without counting.
3. Copies a written model.
4. Copies a diamond.
5. Repetition of five digits.

6. Description of pictures.
7. Counts 13 coins.
8. Knows names of four common coins.

AGE 8

1. Reads a passage and remembers two items
2. Counts nine coins.
3. Names four colours. The four primary colours
4. Counts backward from 20 to 0.
5. Writes from dictation.
6. Differences Comparison of two objects.

AGE 9

1. Knows date, day of week, day of month, name of month and year.
2. Repeats days of week
3. Makes change (with coins).
4. Definitions superior to use.
5. Reads a passage and remembers six items.
6. Arranges five boxes in order of weight.

AGE 10

1. Repeats the months of the year.
 2. Knows the names of nine pieces of money
 3. Uses three words in one sentence.
 4. Comprehension of common-sense questions (easy)
 5. Comprehension of common-sense questions (difficult).
- Thus last test Binet called a transitional test between the ages ten and eleven.

AGE 11

1. Criticizes absurd statements.
2. Uses three words in one sentence.
3. Names sixty words in three minutes.
4. Defines abstract words.
5. Disarranged words arranged into a sentence.

AGE 12

1. Repetition of seven digits.
2. Finds three rhymes for a given word.
3. Repetition of a sentence of twenty-six syllables.
4. Interprets pictures.

AGE 13

1. Paper cutting, described in 1905 scale, test 29.
2. Reversed triangle, rearranges two triangles in imagination and draws results.
3. Differences between pairs of abstract terms.

The construction of this list seems a very simple thing for Binet to have done but it was the result of years of experimentation during the early dawn of psychology as a science.

The significance of this scale of Binet's is that a really usable way of comparing the intelligence of various people was given to us. What one child can do in comparison with other children of the same age, gives the clue. It became clear that a ten-year-old child doing the tests of an eight-year-old child was retarded two years out of ten. A ten-year-old child able to do the tests expected of the normal twelve-year-old child was obviously of the same mental ability or *mental age* as the twelve-year-old. This is the really significant idea that Binet gave to us, that of mental age. We know a great deal about a child when we hear that his chronological age (which means his age calculated in years) is ten and his mental age is also ten. We know immediately that such a child is about normal in mental ability. If the mental age is ten and the chronological age is eight, we know that the child is above normal, and by how much he is above normal. If his mental age is less than his chronological age, we know that he is retarded and if the exact ages are given, we know just how much he is retarded and to some extent what his prospects of development are. Binet gave us, in the concept 'mental age' a way of classifying people according to their native ability or speed of learning. He could now classify his mental defectives and plan their

education more carefully. Incidentally he also taught us to classify our normal and superior children.

Three years later, shortly before his death, Binet improved his scale. Psychologists soon recognized the value of Binet's work. By 1913 experimenters in Belgium, England, the United States and Germany had adapted Binet's tests and enlarged upon them. (In England the much-used version by Cyril Burt, is a rather recent version.)

In America, Lewis Terman benefiting from the experiments in various countries greatly improved the scale, giving us (in 1916) what is known as *The Stanford Revision of the Binet-Simon Scale*. Terman popularized a term which has been found to be quite practical, namely 'intelligence quotient' or, as it is called for short, the IQ. The IQ is the ratio between the individual's mental and chronological age.

$$\frac{MA}{CA} \times 100 = IQ$$

MA means Mental Age; CA means Chronological Age. Thus, if Pyare Lal's mental age is found to be ten and his chronological age is eight, we may write

$$\frac{10}{8} \times 100 = 125$$

The multiplication by one-hundred is done to get rid of the decimal point. An IQ of 100 is just average; 125 is quite high, and 75 denotes a low grade of intelligence. The IQ of a person is assumed to be the same all through life. It has recently been found that the IQ undergoes some change during the course of the years of the individual life but for practical purposes the layman may well assume the IQ to be comparatively stationary throughout life.

After widespread experiment, Terman concluded that the mental age of the average person increases only up to about age 16. This does not mean that the adult does not know more than the sixteen year old youth. Intelligence is not a matter of what we know. An adult's knowledge and maturity are certainly greater than that of the sixteen-year-old but the rate at which we can learn new things is not much greater at 25 or 40 years of age than it was at sixteen years of age. Intelligence means ready adaptability to new situations, alertness, keenness, ingenuity. It is quite different from knowledge and experience. We by no means reach our maximum powers at sixteen (we reach these powers nearer fifty if we avoid mental laziness); but the normal youth, if placed in a situation in which previous experience gives no advantage, has a chance to meet the situation just as intelligently as the normal adult.¹

Practical use of Intelligence tests on a large scale was first attempted by the United States Army after the entry of the United States into the World War. At the time of

¹ Popularly the word 'intelligent' is used to mean a variety of qualities, such as clever, clear-headed, well-informed, sensible, shrewd, quick-witted, etc. The mental tester uses the word in a much narrower sense; he means by an intelligent person, a person who scores well on an Intelligence test, or, in other words, has a high IQ. When he says that A is more intelligent than B, he means that A secured or is expected to secure a higher score in an Intelligence test. He ordinarily expects A also to show more of the kind of intelligence which means common sense and cleverness in everyday life but we cannot guarantee that this will be the case. He will not be at all surprised if occasionally the opposite is the case, for it takes more than IQ to make common sense and shrewdness in everyday affairs, although the person with a high IQ has decidedly the advantage in developing these.

The reader must not confuse these two different meanings of 'intelligent'.

its declaration of war against Germany in 1917, the United States had an army of only about 100,000 men. The task was to train an army of several million men in less than two years. In order that much time should not be wasted on those of too low intelligence for duty and in order to pick out the men who could be expected to adapt themselves most readily to positions of intelligent responsibility, psychologists were invited to try their intelligence-testing technique. By this time a number of other scales besides that of Binet were ready or almost ready for use. Binet's scale did not meet the requirements as it has to be given to each individual separately. Tests were needed that could be given to large numbers of people at one time. A. S. Otis and R. M. Yerkes, who had prepared tests of this type, were among the most prominent of the psychologists entrusted with the task of rapidly testing hundreds of thousands of men.¹

One set of tests was made for men of education. This is still called *The Army Alpha Test*. The set was comprised of ten alternate forms complete in themselves, each form consisting of eight parts, the large number of items in each part making coaching practically impossible.

Another set of tests was made for those who had little or no school education and for foreigners who did not understand English well enough to take a test such as the Alpha. This was called *The Army Beta Test*.

Some two million men were tested and classified according to results on these tests. The results of the use of the tests were valuable beyond all expectations. It was proved that classification of men on this basis for the kind of work they were to do was far more satisfactory than on

¹ A lengthy account of the experiments is given by Yerkes in his book, *Psychological Examining in the United States Army*. Government Printing Office, Washington, 1921.

any other basis of classification hitherto suggested. The army Intelligence tests did not try to pick out men for specialized jobs such as infantrymen, cavalrymen, clerks or gunners. The army tests simply picked out men for responsible positions on the basis of rapid adjustability as exhibited in their performance of the tests. The outcome was a greatly increased respect for the psychologist and for intelligence-testing.

Actually there was too much respect for intelligence-testing. Some very exaggerated claims were made for them by psychologists and especially by would-be psychologists. Describing the advent of the Army Tests, Goodwin Watson writes:

The epoch of high hopes and exaggerated expectations was set off with a bang by the World War. Intelligence tests were expected to identify not only the feeble-minded but also the talented, to predict school success and also vocational success, to prognosticate delinquency, to reveal native endowment, and to answer scientifically long disputed questions of race difference and sex difference.¹

After the period of 'exaggerated expectations' came the period of 'soberization' in which serious hard work was needed (and was done) to discover the limitations and legitimate sphere of intelligence-testing. Also a good deal of light was wanted on the question: 'What is intelligence? This question has proved to be almost as elusive as the question 'What is electricity?' One of the major controversies of psychology sprang out of the attempt to answer this question.

¹ 'The Scientific Movement in Education' in *The Thirty-seventh Year Book of the National Society for the Study of Education*, Part II, ch. XXIV, p. 358, 1938. Published for the National Society for the Study of Education by the Public School Publishing Company.

Not much headway could have been made in understanding the benefits and limitations of intelligence testing if about this time the method of handling and relating data had not been greatly improved. Here is England's unique contribution to psychological testing; the refinement of statistical treatment. Galton, Pearson, Brown, Thomson and Spearman perfected the machinery by which the data obtained by experimentation could be properly treated and brought to bear on the problems affecting intelligence.¹

What is intelligence? How far is it influenced by education? By environment? To what extent does our original mental inheritance limit the amount of our intelligence for ever? Do Intelligence tests really test intelligence or something else? What kinds of intelligence? Which Intelligence tests are useful and which are not? Just what do various tests show and what do they not show? All these questions need to be answered before Intelligence tests become really helpful and all of them are related to the question: 'What is intelligence?'

Binet held that intelligence is made up of three important abilities:

1. The ability to comprehend a problem, direct the mind towards its solution, and maintain attention upon it.
2. The ability to adapt the mind to the needs of the situation.
3. The ability for self-criticism.

Quite a debate arose among the psychologists as to whether intelligence is a single complex quality that expresses itself in various ways, or whether it is a large number of independent traits. Among those who challenged the

¹ Chapter X deals in a very elementary way with the treatment of data.

view that it is a unity were Thorndike and Woodworth of Columbia University. These two men gave a blow to the traditional view of educationists which assumes that 'by hard study and application the fibre of the mind becomes toughened like a muscle, and that, as a direct result, one's powers of attention, memory, reasoning and the like are markedly strengthened and increased,'¹ from which it has never recovered. They proved, in a series of brilliant experiments, that the 'transfer of training' (for example, the effect of learning one poem on the learning of another, or the learning of Latin upon the learning of other school subjects) is far less than educationists supposed. This led to the statement of the view that learning is not general but specific and that the mind is composed of a host of highly particularized and independent faculties. These views seriously effect our view of what intelligence is, for if the mind is composed of particularized and independent faculties it is not unreasonable to suppose that intelligence also is made up of a large number of independent traits. Are there not perhaps many different kinds of intelligence? We know that some people are 'intelligent' in some things and 'stupid' in others. The same individual is not alike intelligent in handling various kinds of situations. The view has been put forth that instead of there being such a quality as 'general intelligence' there is simply the average of our various kinds of intelligence.

The experiments of Thorndike have been of immense practical value in education and in psychology. Perhaps

¹ This is how H. E. Garrett in *Great Experiments in Psychology*, p. 127 (The Century Co., New York, 1930) describes the doctrine of formal discipline which is the traditional view of education. The formal disciplinists have held that the study of mathematics and classical languages increases the acuity of the mind much more proportionately than other mental activities do.

no discussion has been more helpful in jarring educational practice out of its worst ruts as the discussion that followed on Thorndike's and his associates' experiments.

But is there no unity in the various intellectual abilities? Is there no general factor underlying the various abilities? Burt insists there are. He claims that intelligence is an 'all round efficiency'. Thorndike also searches for a 'unitary factor', although he has demonstrated the specific nature of abilities and learnings.

Spearman provides for both the specific and the general nature of intelligence in his theory. He recognizes both a 'sum total of mental energy' which is present in all our mental efforts and on the other hand specific abilities. The first he calls 'G' and the second 'S'. His 'G' is akin to 'general intelligence' and his 'S' to 'specific intelligence', but he rightly avoids the word 'intelligence' because the many meanings popularly associated with this word are misleading. An individual may have many 'S's but only one 'G'.

The original data submitted in support of this theory are mathematical. By means of the tetrad equation (this is a highly technical device the description of which must be left to works of a more technical nature than this book purports to be) he proves the existence of both 'G' and 'S' and the extent to which various abilities overlap. He then proceeds to prove his theory experimentally also. He has made comprehensive studies of the relationship of various abilities with a large number of other abilities which bear out his theory. Spearman's theory has done much to answer the question 'What is Intelligence?' His theory is generally credited with being nearest to a satisfactory explanation.

Meanwhile experimentation has gone on. We still need to learn a lot about what intelligence is and what it is not,

but we can undoubtedly measure it at least roughly. In place of the half-dozen Intelligence tests of twenty years ago there are now literally hundreds with which to test everyone from infants to adults. There is hardly an aspect of education that has not been affected by them. Especially the treatment of delinquents and of the mentally inferior and superior, bifurcation of courses, individualization of instruction, personnel guidance, and mental hygiene have been greatly benefited.¹ Ballard, in his usual common-sense way, gives an excellent lead as to how the teacher may profitably regard them

It would be idle, therefore, to maintain that mental tests as they exist today finally settle any of the theoretical issues of psychology or education. They do not succeed in dissecting the mind so as to separate innate abilities from acquired abilities, or to separate general ability from specific abilities, or to separate mother-wit from book-learning. Ideal mental tests would do these things; actual mental tests, if they do them at all, do them very roughly. This every mental-tester is ready to acknowledge, and he does not care a dump if the whole fabric of biological and psychological theory that has been prematurely built up on mental tests comes crashing to the ground. The mental-tester does not care because, whatever happens to the superimposed theory, mental tests remain as they are—an inestimable boon to the teacher who wishes to discover the educability of his scholar.²

¹ An excellent appraisal of what Intelligence and Aptitude tests have contributed to education is Chapter XXX of the *Thirty-seventh Year Book of the National Society for the Study of Education* from which the quotation (p. 142) by Watson is taken. The full contribution of intelligence testing is far greater than is indicated in the outline of uses of such tests given on pp. 147-150 of this book, for the uses described are meant for the ordinary teacher and not the expert.

² *The New Examiner*, p. 121.

A conservative psychologist will never decide the future of a person on the basis of one Intelligence test alone. To him any information he can get as to what the person examined can do and how he behaves in ordinary life is valuable supplementary information. If possible, he gets a measurement on two or three Intelligence tests; particularly if the information he gets as to the person's ability otherwise does not support the information he got from the test. For instance, if a certain boy according to a test has a low IQ but is doing good work in school without being over age for his class, and gives the impression of being alert, one may well question whether he did himself full justice in the Intelligence test. It would be well to test him again, if possible on a different type of test. An Intelligence test is valid only if the examinee did his best in that test. One can easily be too cocksure of the results of Intelligence tests but nevertheless they do give us valuable information that is not obtained nearly so well in any other way.

On the use of Intelligence tests, an outline of possible uses given by Dr R. A. C. Oliver in the *Manual of Directions of his General Intelligence Test for Africans* applies very well in the case of India also.¹

1. *The Selection of Entrants to a School*

There are often more applicants for admission to a central or secondary school than can be accommodated. Yet all of the applicants may have reached the educational standard required for admission. It then becomes necessary to make use of some additional means of selection. In that case the use of an Intelligence test as the selective instrument may be considered. It will have the advantage of eliminating the influence of some of the more adventitious circumstances which may have

¹ Published by the Government Printer, Kenya Colony, 1932 and reprinted here by kind permission of the author.

contributed to the gaining of the necessary school certificates. For example one of the applicants may have reached the required standard by dint of hard 'cramming', or of specially good teaching or simply of 'easy marking' of his examination papers; while another may have qualified for entrance in spite of poor teaching or of having to herd goats too often when he should have been at school. The Intelligence test would pick out the innately brighter boy.

2. *Vocational Guidance*

A central school may provide a variety of courses of study, such as a general course, a commercial course, a teacher's course and a handwork course. The principal of the school may be called upon to direct a pupil into one or other of these courses. One of the items of information which might then be useful to him would be the pupil's score in an Intelligence test. Suppose that the principal knew, for example, that his teachers' course called as a rule for a higher degree of general ability than his handwork course, he would, while taking other circumstances into consideration, tend to advise the teachers' course for his brighter boys.

3. *The Classification of Pupils*

Pupils vary enormously in intelligence, or the ability to learn. In almost any class of Standard III pupils, for example, there will be found some pupils who are brighter than the average pupil in Standard V, and others who are not so bright as the average pupil in Standard I. To teach fractions, say, to children at such different levels of mental development, by the same method and at the same rate, is obviously a difficult and certainly a wasteful task. Now in a school with a large enrolment it is sometimes necessary to divide a standard into two or more sections. Where this is so, it is advantageous both to teacher and pupils if the pupils are grouped on the basis of learning capacity; and of learning capacity an Intelligence test score is a fairly reliable index. There should be three sections, if possible—one for the backward pupils, one for the

advanced pupils, and one for those round about the average. Each section will be able to progress at its own rate. The dull pupil will not be discouraged by a hopeless competition against his intellectual superiors. The clever pupil will be able to finish his course more quickly, or to cover much more ground in the same time. The teacher's task will be greatly simplified. Even if it is not possible to run three separate classes, it may be possible to group the pupils within one class according to ability, for some subjects at least; or else the more advanced pupils may be moved up to a higher class sooner than the others. In one way or another the instruction should be adapted to the individual differences in learning capacity revealed by the Intelligence test.

4. *The Discovery of Unusual Cases*

In the process of applying an Intelligence test to a class he does not yet know, the teacher may discover, sooner than he otherwise would, pupils who suffer from some mental or physical abnormality. A pupil whose score is greatly below the scores of other pupils whom he might have been expected to equal, may quite possibly be found to suffer from defective vision or hearing, or from some maladjustment of personality.

An Intelligence test will often throw new light on pupils whom the teacher thought he already knew quite well. He will probably be surprised at the range of ability in his class, for one's almost inevitable tendency is to overestimate the intelligence of the duller pupils and to underestimate that of the brighter. He may come to realize the need for special opportunities for his best pupils, who, after all, are to be the leaders of their people. He will be reconciled to the slow progress of most of the weaker brethren, realizing that they may be doing very creditable work in view of their limited abilities. On the other hand, it may come as a shock to him to find that a pupil whom he had definitely relegated to the ranks of these same weaker brethren may turn out to be one of the most intelligent

boys in his class, and he will take steps to ensure that his talent is no longer buried.

5. *Vocational Selection*

The employer's task in selecting a candidate for a post where intelligence is a desideratum, is analogous to that of the school principal in selecting applicants for admission to his school. An Intelligence test will indicate which candidate is likely to adapt himself best to the new conditions of work; and if he is qualified in other respects, this is the candidate whom the employer will choose.

6. *Research*

Scientific research requires the use of accurate measuring instruments. A good Intelligence test is such an instrument, and the use of one in East Africa would begin to make practicable a large number of important and fascinating researches. Some of the questions an Intelligence test might be instrumental in answering are as follows. How do the East African tribes compare with one another in intelligence? Are native women less intelligent than native men? At what age does the average African's intelligence cease to develop? At what period of old age does it definitely decline? Is there any relationship between intelligence and stature in the native? Is the size or shape of his head a clue to his intelligence? Is the African's intelligence chronically depressed by diseases, and, if so, which diseases have this effect? Is any rise in the general level of intelligence visible over a period of years?

It is utterly impossible to give an idea of the contents of more than a very few Intelligence tests, nor is there much object in trying to describe them, for an adequate notion of a mental test can be gained only by study of the complete test. Description of particular tests will be confined to those which have been adapted for use in India.

Many of the more prominent Intelligence tests in England are included in Cyril Burt's book *Mental and Scholastic Tests* and in P. B. Ballard's *Group Tests of Intelligence*, *The New Examiner*, and *Mental Tests*.¹

Other prominent tests are:

<i>Group Intelligence Tests</i>	by Godfrey Thomson
<i>Simplex Intelligence Test</i>	by C. A. Richardson
<i>Northumberland Intelligence Test</i>	by Cyril Burt
<i>A Measure of Intelligence</i>	by C. Spearman
<i>Cattell Group Intelligence Tests</i>	by R. B. Cattell
<i>Psychological Tests of Educable Capacity</i>	Board of Education, England

Pintner lists the following thirty-seven group tests as the most commonly used and readily available in America.²

- | | |
|--|---------------------------------------|
| 1. <i>Army Alpha</i> | Stoelting Company,
Chicago. |
| 2. <i>Army Beta</i> | " |
| 3. <i>CAVD Scale</i> | T. C. Bureau of Publications,
N.Y. |
| 4. <i>Detroit First Grade</i> | World Book Co. |
| 5. <i>Detroit Advanced First Grade</i> | " |
| 6. <i>Detroit First Grade</i> | Public School Publishing Co. |
| 7. <i>Detroit Advanced</i> | " |
| 8. <i>Dearborn, Series I</i> | Lippincott, Philadelphia |
| 9. <i>Dearborn, Series II</i> | " |

¹ For a list and appraisal of more recent intelligence tests in England and to a less extent in America see the *Guide to Mental Testing* by R. B. Cattell, University of London Press, 1936. Anyone wishing to get acquainted with the different types of tests and the principles involved in construction of different types will find this reference the handiest one.

² Pintner, R. H., *Intelligence Testing*, p. 216. Holt & Co., New York, 1923.

- | | |
|---|-----------------------------------|
| 10. <i>Haggerty, Delta I</i> | World Book Co. |
| 11. <i>Haggerty, Delta II</i> | " |
| 12. <i>Illinois Examination</i> | Public School Publishing Co. |
| 13. <i>Kingsbury Primary</i> | " |
| 14. <i>Kuhlmann-Anderson</i> | Test Bureau, Minneapolis. |
| 15. <i>McCall Multi-Mental</i> | T. C. Bureau of Publications. |
| 16. <i>Miller Mental Test</i> | World Book Co. |
| 17. <i>Myers Mental Measure</i> | Newson & Co., New York. |
| 18. <i>National Intelligence Test</i> | World Book Co. |
| 19. <i>Ohio State University Test</i> | Ohio State University, Columbus |
| 20. <i>Otis Primary Test</i> | World Book Co. |
| 21. <i>Otis Advanced Test</i> | " |
| 22. <i>Otis Self-Administering Tests</i> | " |
| 23. <i>Otis Self-Classification Tests</i> | " |
| 24. <i>Pintner-Cunningham Primary</i> | " |
| 25. <i>Pintner Primary Non-Language</i> | T. C. Bureau of Publications. |
| 26. <i>Pintner Non-Language Mental</i> | College Book Co., Columbus, Ohio. |
| 27. <i>Pintner Non-Language and Educational</i> | " |
| 28. <i>Pintner Rapid Survey</i> | T. C. Bureau of Publications. |
| 29. <i>Pressey Primary</i> | Public School Publishing Co. |
| 30. <i>Pressey Cross-Out</i> | " |
| 31. <i>Princeton International Tests</i> | Princeton University. |
| 32. <i>Revised Army Alpha</i> | Psychological Corporation, N.Y. |
| 33. <i>Rhode Island Intelligence</i> | Public School Publishing Co. |
| 34. <i>Ternian Group Test</i> | World Book Co. |

35. *Thorndike Intelligence Test* T. C. Bureau of Publications.
36. *Thurstone Psychological Examination* Stoelting Company.
- 37 *Trabue Mentimeters* Doubleday, Page, N.Y.

VII

INTELLIGENCE TESTS IN USE IN INDIA

THE problem involved in adapting an Intelligence test for use in India is quite different from the problem involved in adapting an Achievement test. Mental tests are very much harder to adapt or create and validate than Achievement tests. It is necessary to understand the difficulties in transferring Intelligence tests from one environment to another before describing the tests now available.

Intelligence tests may be divided into three classes according to the extent to which they employ language. (i) non-linguistic; (ii) non-verbal; (iii) verbal.

Non-linguistic tests make no use of language whatever, not even in giving directions to the examinee. All directions are given by means of charts, pantomime. There are not many tests in this class, *The Army Ohio Test* being one of the few. They are used mainly for people who do not understand the language of the examiner and for the deaf. In tests for very small children and backward children some such items are usually included.

In non-verbal tests there may be any amount of speaking on the part of the examiner, but very little or no use of written language is required of the examinee. All tests for children or adults who cannot be depended upon to read well must be either non-linguistic or non-verbal. Adapting a non-verbal test in a foreign country is not nearly as difficult as adapting a test in which language mastery is required by the examinee. In a country of

many languages, like India, the non-verbal tests are much to be preferred, especially since so many pupils must be examined in a language not their actual mother-tongue and many of the actual mother-tongues are either not reduced to script or have an almost negligible literature.

Verbal tests require reading and writing skill. A scrutiny of verbal tests will soon reveal how essential a thorough knowledge of the language is. Tests using word-analogies, synonyms, opposites, straightening out sentences in which the words have been mixed up, etc., are quite common. In items testing reasoning and judgement considerable reading skill is required. Such tests are very difficult to translate and it cannot be assumed that the translation is equivalent to the original in difficulty. There may be, for all we know, quite a different degree of intelligence and skill required in the reading of Urdu, Hindi and English. The words used in analogy, synonym, and similar tests may be less common in one language than their equivalents in the other. Translation makes a different test, to which the norms of the original test are not applicable.

In India pupils of classes from Class I to college rank do on the whole less voluntary reading than students in equivalent classes in England or America. From the experience gained in standardizing reading tests both in primary and in middle schools (Hindi), the writer is convinced that even in Classes VII and VIII at least 35 per cent of the students read so poorly as to be severely handicapped in any mental test requiring more than the simplest kind of reading. It is certain that the ability to read varies so much between student and student coming out of the primary school, that any mental test used for purposes of choosing students for admission to middle school, which contains reading exercises of any difficulty,

is bound to be a Reading test as much as an Intelligence test and pupils will be chosen on the basis of reading ability as much as on the basis of intelligence. In general we may say that the best readers will also be the most intelligent, but there are many exceptions and the rule holds true only if the opportunities of learning reading are about the same, and there are no special disabilities or defective habits. Reading is taught so much more poorly in some primary and middle schools than in others, that the incompetence of a teacher will do injustice to many a child whose intelligence is erroneously estimated on the basis of reading ability. The verbal type of mental test should be used with great caution below high school standard, and even in high school reading ability can easily be overestimated. The most practical type of Intelligence test for either children or adults who have not had at least some high school experience is the non-verbal, although the verbal can be used if the examinees undoubtedly have had both adequate teaching in reading and adequate reading facilities, the fulfilment of which conditions are all too rare in India.

✓ An indication of intelligence in a child is present when he learns a great many things from his environment on his own initiative without any one having taken any pains to teach them to him. The things that an Indian child will learn of his own accord from his environment are quite different from the things an English or American child will learn. In England it may be assumed that any child of nine years of age will have learned to tell the time even though this has not been taught in school. He sees so many clocks and everyone lives so much 'by the clock' that he finds it necessary to be able to tell the time and insists on being taught. In an Indian village there may not even be a clock (even some schools are

without clocks) and village life pays little attention to being 'on time'. Thus lack of ability in a child to tell the time does not reveal lack of intelligence in India in the way it does in England. In America, where there is one motor car to every four people, any intelligent child of eight knows something of motors and simple mechanics. Any item in an American Intelligence test that bears on this field of experience may be fair to an American child but certainly not to an Indian. The Indian child, on the other hand, learns a great many things from his environment that an American or English child does not. Just what these things are, no one has taken the trouble to record as yet, but there will be such items. In Africa Dr Oliver found the 'Maze tests', which are composed of a series of ground-floor plans of labyrinths with several blind alleys and only one way out, which constitute quite an exacting test for American children, were done so easily by the cautious African children who are used to finding their way through jungle that they were too easy for them.

It has been found that, on the average, the people of some races and nationalities do less well in certain Intelligence tests than the people of other races and nationalities. This has led some people to claim that the people of their own race or nationality are on the average more intelligent than the people of certain other races. Any such interpretation of the results of Intelligence tests is based on a number of false assumptions. There is no Intelligence test in existence that can show Indian children on the whole to be either superior or inferior in mental ability to the Chinese, French, English, Negroes, or Americans. Intelligence tests now available can compare Indian children and American children only if the environment, education, and culture of the two countries are nearly

alike, which they evidently are not. If we sent several thousand Indian children to be adopted by American parents living in America, and several thousand Americans to be adopted by Indian parents in India, we might be in a position to compare the relative intelligence of the two peoples, but not at present. The fact that Indian children do not compare well with American children on American Intelligence tests may mean either that the American children have superior cultural, educational, social, and economic advantages, or that the tests give the advantage to those who live in the American cultural environment. There may be tests developed some day at which Indian children will do much better than the people of any other environment.

Binet insisted that his scale could be used only on people living in a fairly homogeneous environment, and that as soon as it was applied to those of a different country, the standards or norms could no longer be considered valid. Binet's and Simon's original scale is used only in France and seldom in America or England. The reason is more than just language differences. In the latter countries it has not only been translated but re-adapted. The English and American scales, though in the same language, are not interchangeable.

It thus follows that any test for India must be adapted to Indian subjects and only Indian standards can be considered valid. Comparisons with other countries may be made but these have no real significance and prove little. The reasons why Indians score either higher or lower than the inhabitants of some other country may be many besides that of superior or inferior native ability. We try to measure native ability but we never quite isolate native ability from environmental effects and as long as these effects vary as considerably as they do now, we

cannot be sure that the difference which appears to be a difference in the intelligence of a race or nationality, may not be something quite different.

The first mental test standardized for India was, quite appropriately, Binet's scale.¹ The honour for this pioneer piece of work goes to Dr C. Herbert Rice, then of Forman Christian College. This work was begun in 1922 and is described by Dr Rice in his book *The Hindustani Binet-Performance Point Scale, Examiner's Manual* ²

Items from several other tests were experimented on at the same time and the final scale as presented is somewhat different from the Stanford-Binet scale which served as a starting-point. Altogether sixty-seven items were tried out. Of these thirty-two were eliminated as being unsuited to Indian conditions, difficult to translate, or not well suited to the age-ranges tested. The languages used were Urdu and Punjabi. Only very slight changes are needed for use in Hindi.

The test items included are:

- Knox cubes (non-verbal)
- Comparing weights (non-verbal)
- Goddard Form-board (non-verbal)
- Repeating digits forward
- Repeating digits backward
- Repeating syllables
- Description and interpretation of pictures
- Questions
- Words in two minutes
- Drawing two designs (non-verbal)

¹ Experiments with the Binet scale were conducted about the same time in the Government Training College, Saidapet, Madras. The results of these experiments are given in *Bulletin No. 15* of the Teachers' College. The Saidapet experiment is not as complete as Dr Rice's.

² Oxford University Press, 1929.

Copying a square
 Copying a diamond (non-verbal)
 Choosing prettier faces
 Recognizing missing features
 Giving number of fingers
 Naming six coins
 Concrete definitions
 Abstract definitions
 Adding 3 two-anna and 3 onc-anna pieces
 Counting backward from 20 to 1
 Making change
 Adaptation board (non-verbal)
 Divided oblong card (non-verbal)
 Naming colours
 Giving differences between concrete objects
 Giving similarities between two things
 Giving similarities between three things
 Sentence building with three words
 Ball and field
 Giving three rhymes
 Detecting absurdities
 Folded paper Test
 Healy Form-board (non-verbal)
 Enclosed boxes
 Raja and deputy commissioner

Norms are given for 929 boys. Later the results of 459 more tests were added.

The scale is conveniently divided into a brief scale of ten items and a longer scale of thirty-five. The brief scale compares well enough with the longer one so that where the examiner's time is limited, the brief scale may be considered satisfactory though not as dependable as the longer scale. The non-verbal items have separate norms and can thus be used separately.

The intelligence quotients obtained by using the Hindustani Binet scale are by no means comparable to those Terman reports on the Stanford-Binet. One need only

examine carefully the following table from Rice's book¹ to see that the range of IQs on the Hindustani is much greater. Therefore the term IQ must be used very carefully in India. An IQ of 130 on an Indian scale is not nearly so hopeful as the same IQ on an American scale nor is an IQ of 65 nearly so hopeless as the same IQ in some other countries. Probably it would be safer to drop the term IQ in India until we know better what significance these figures have. Age and grade norms are quite usable even without translation into IQs.

PER CENT OF CASES	CLASSIFICATION	STANFORD IQ	HINDUSTANI IQ
Highest 0.5%	Genius	140 and up	165 and up
Next 5.0 "	Very Superior	120-140	140-165
Next 15.0 "	Superior	110-120	120-140
Middle 60.0 "	Average	90-110	85-120
Next 15.0 "	Dull	60-90	70-85
Next 5.0 "	Border-Line	70-80	55-70
Lowest 1.0 "	Feeble-Minded	Below 70	Below 55

V. V. Kamat has developed a Binet Scale in Marathi and Kannada (*Teaching*, September 1935) which is described fully in Mr Kamat's books in both these languages and also in English entitled *Intelligence Tests for Indian Children*.² 1074 children were examined. Kamat's scale is different from Rice's in a number of details. He too found that the range of IQs is much greater in India than Terman reported for America. Kamat did a creditable and painstaking piece of work.

Kamat reports better scores for boys than for girls.

¹ *A Hindustani Binet-Performance Scale*, p. 122. Oxford University Press, 1929.

² Oxford University Press. English edition published in 1940.

Rice examined only boys so that no comparison can be made.

Several more revisions of Binet tests have recently come to the attention of the author. These are associated with the institution in which they were applied.

Dacca Training College ...	<i>Pal's Stanford Revision in Bengali.</i>
Patna Training College ...	<i>Stanford Hindustani Revision.</i>
Shikarpore, Bombay ...	<i>See Education Pamphlet No 28, Government of India.</i>
Lady Willingdon Training College, Madras	<i>Stanford Revision in Tamil and Telugu.</i>
Khajua, U. P.	<i>Gupta's Binet Tests in Hindi.</i>
Calcutta University	<i>Maiti's Adaptation of the Stanford Revision (Bengali).</i>

The Binet scale is traditionally considered the most valid of all Intelligence tests, and its results, when reported by a well-trained examiner, the most trustworthy of all. Any test that does not correlate highly with the Binet is rightly viewed with suspicion although the Binet scale is by no means infallible and there are important elements of intelligence that more recent tests examine better.

A decided limitation of all Binet scales is that it takes a well trained examiner to administer them. Any one can try his hand at administering them but the details of procedure are so many and the effects of slight changes in the way of administering the various tests so considerable that psychologists warn the novice from trying to administer a Binet test until he is thoroughly grounded in psychology and has spent some months under an

expert. The other limitation is that the Binet scale is administered to one person at a time, each test taking about one hour. It is not surprising that group tests are more popular than this more trustworthy but time-consuming scale which can be entrusted only to an examiner who has had the advantage of careful training.

As late as 1938 a new test has come into use called the *Lester International Performance Scale* which may become a step in advance of the Binet tests for use among more primitive people and for the comparison of children from highly diverse environments—from the most cultured to the most primitive. It is an individual non-verbal performance test, testing from the two-year old to the adult level. Unfortunately it requires a very elaborate apparatus. Mrs E. Earle of Nowgong, C I., has been experimenting with it in India but has no results to announce as yet. It will repay psychologists in India to give attention to this test which is the result of a long research to find test items which are less dependent on a certain uniform cultural level of examinees than the Binet test is. I have not had the opportunity of seeing the apparatus.

VERBAL GROUP TESTS

The first verbal group test adapted to Indian conditions, it seems, is the *Preliminary Classification Test* of Dr J Manry of Ewing Christian College.¹ This was published in 1927 in Urdu, Hindi and English.

Its author asserts that

It is not intended to take the place of the individual intelligence examination. Nor is it to be used outside of schools or below Standard II. The advantage

¹ Published by the North Indian Tract and Book Society, 18 Clive Road, Allahabad.

of the test is that a general idea much more accurate than the teacher's casual impressions, though still fallible, can be obtained of the ability of a large number of children in a short time.

This is a combination of a Scholastic and an Intelligence test. As its title claims, it is meant to be useful for a tentative classification of pupils. The hundred questions of the test alternate in rotation: Information, Judgement, Language, Number. Age-norms are supplied but not class placement norms.

In 1933 Pundit Lajja Shankar Jha published in the *Teachers' Training College Annual*¹ of Benares Hindu University a report of the examination of over 1,000 pupils on a Hindi adaptation of the *Simplex Mental Test* by C. A. Richardson. This test contains items on:

1. Ability to pick out words not in their class
2. Ability to give reasonable answers to specific questions.
3. Ability to identify indispensable parts of an animal or object
4. Ability to see similar relationships.
5. Ability to grasp accurately what is told and to act accordingly
6. Solving sentence tangles.
7. Seeing arithmetical relationships.
8. Ability to understand nice shades in the meanings of the words.
9. Filling in gaps of thought.
10. Ability to visualize a concrete experience with certain modifications.

It is intended for children between the ages of 10 and 18. Pundit Jha complains of linguistic difficulties on the part of the younger examinees. He thinks that the intro-

¹ A revised report was printed in *Science and Culture*, February - May 1939.

duction of English into the school from Standard I retards the linguistic development of the mother-tongue. The remarks made on the difficulty of tests demanding considerable reading ability on the part of the examinee probably pertain here.

Pundit Jha also adapted Terman's *Group Test of Mental Ability*.¹ This test is intended for the same ages as the *Simplex* and seems, from a superficial examination, to contain more appropriate material for the younger children of this age-range than the *Simplex*.

S. Jalota of the D. A. V. College, Lahore has prepared a group verbal test for use among college students.² It is prepared in Hindi, Urdu, and English. As this test is typical of the style of test generally used in high school and college a sample of the various tests comprising the battery is given here.

TEST I (FOLLOWING DIRECTIONS)

Directions

Below are twelve instructions; you are to do what you are directed to do as quickly as you can. When I say 'Stop', stop instantly at what you are doing and hold your pencil up.

1. Write ten in figures in the largest square.



2. Ram has four big balls:—white, red, green and blue. He has given the green one to Hari, the white and the blue ones to Sita. Underline which of the following he has kept?

white red green blue

¹ Both tests are obtainable from the Teachers' Training College, Benares Hindu University

² *Teaching*, September 1936 and March 1937.

3. Underline the number that follows next but one after 19.

17

21

22

23

TEST II (NUMBER SERIES)

Directions

Look at each row of numbers below, find out how the numbers are made up, and then on the two dotted lines write the two numbers that should come next.

1.	8	7	6	5	4	3
2.	3	6	9	12	15	18
3.	9	9	7	7	5	5
4.	1	2	4	8	16	32
5.	8	9	12	13	16	17

TEST III (MEMORY)

Directions

Read this paragraph to yourself. Later, you will be asked to recognize some of the words used here. So read it with great care.

How is one to judge change within a religion as complex as Hinduism? The answer seems to be that it will be found chiefly in terms of social change, since the one common element that binds the Hindus together is not belief but practice of certain social customs; in their own phrase, 'the Dharma'. Religious ideas are unquestionably in a fluid state in India and we shall show some evidence of the change, but the vastly more important changes are to be found in the realm of social organization and practice. Here in the modern age, which is causing such revolution in the religious thinking of other peoples, is having its deepest effect. That thought and practice are bound up together none will seriously question, and that corresponding changes within the thought life of particular groups in India will come about as a result of accepted change in social organization and practice seems clear.

TEST III (B)

Below is a list of words. Some of them have been used in the paragraph you have just finished reading. Some other words in this list have not been used.

In front of each of the following words are written the words 'yes-no'. Now, if you think that the given word has been used in the paragraph that you read, draw a line under YES, thus 'yes-no'. But if you think that the given word has not been used there, you underline NO, thus 'yes-no'.

- | | |
|------------|--------|
| 1. Reply | yes-no |
| 2. compare | yes-no |
| 3. complex | yes-no |
| 4. link | yes-no |
| 5. change | yes-no |

(There are 20 such words)

TEST IV (CLASSIFICATION TEST)

Examples

chair	table	bed	stove
Japan	China	India	France

In each line cross out the word that does not belong there. Cross out JUST ONE WORD in each line.

- | | | | | |
|--------------|---------|-------|-----------|--------|
| 1. Rama | Sita | Shyam | Prem | |
| 2. ears | eyes | heart | nose | tongue |
| 3. motor car | bicycle | tonga | telegraph | train |
| 4. baby | doll | clad | kid | kitten |

TEST V (JUDGEMENT)

Read each question or statement, and make a cross before the BEST ANSWER.

1. Cats are useful animals, because
 1. They catch mice
 2. They are gentle.
 3. They are afraid of dogs.

2. Why is leather used for shoes? Because
 - 1 It is produced in all countries.
 - 2 It wears well.
 3. It is an animal product.
3. Why judge a man by what he does rather than by what he says? Because
 1. What a man does shows what he really is.
 2. It is wrong to tell a lie
 3. A deaf man cannot hear what is said.

TEST VI (ANALOGIES)

Samples:

Sky is to Blue as Grass is to
table green warm big

Fish is to Swims as Man is to
paper lime walks girl

In each of the lines below the first two words are related to each other in some way. What you are to do in each line is to see what the relation is between the first two words, and underline the word in italics that is related in the same way to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

- 1 January is to February as First is to
March *third* second April
2. Moon is to Earth as Earth is to
sea *fish* Mars sun
3. Day is to Midday as Night is to
dark sleep midnight bed

TEST VII (CANCELLATION OF LETTERS' TEST)

Sample:

'We should always honour brave men and women.'
 Every 'n', 'r', and 's' has been crossed out.

A small sentence follows for practice. Try to cross out all those letters in this sentence that have been crossed out in the sample. Remember to cross every 'n', 'r', and 's'.

'One morning there was a traveller in the woods near Shahdara, in the summer season'.

TEST VII (B)

Be careful and quick. Cross out every 'n', 'r', and 's'.

'Asoka is esteemed by many scholars to have been the noblest monarch in history, and if the criterion is the number of souls that still revere his memory, then certainly he was a far greater figure than any other in the whole world's catalogue of kings. By acquiring one state after another he built up an empire that included a large part of the East; and every inch of it he won by faith and not by sword. Asoka sent out Buddhist missionaries to Ceylon, to Kashmir, and to the uttermost ends of the earth known to him. Eight and twenty years he carried on his far-flung missionary work, and before he died he had managed to make Buddhism the dominant religion in his half of the world.'

TEST VIII

In each sentence draw a line under the two words that tell what the thing always has; as in the sample;

Sample:—A man always has a

body *cap* *gloves* mouth *money*

1. A snake always has
poison *rattle* *stripes* *tail* *tongue*
2. A horse always has
harness *hoofs* *shoes* *stable* *tail*
3. A circle always has
altitude *circumference* *latitude* *longitude*
radius.

TEST IX (INFERENCES)

This is a test of reasoning. Overleaf some facts are given; and a question is asked about these facts. Certain answers (in italics) are also given. You are carefully to consider the given facts and then draw a line under the right answer, as in the sample:

Examples

Three boys are sitting in a row;
 Rama is to the left of Abdulla
 Krishna is to the left of Rama.

Which boy is in the middle? *Rama Abdulla Krishna*
 Here Rama is the right answer, so it is underlined

The Pasrur Group Intelligence Test was prepared by R. R. Kumria of the Central Training College, Lahore. The tests in this battery consist of choosing the best answer, absurdities, synthetic wholes, completing number series, Multiple Choice, and reasoning. It was designed for boys entering the upper classes of high school and intermediate college. It is in the Urdu language. Full standardization of the test has not yet been completed but may be expected soon.

The Group Test of Mental Ability for College Students by Dr Lalit Kumar Shah of Lucknow Training College is similar in plan to the two last-mentioned.

Other verbal group tests that have lately been brought to public attention are

The Perera Visual Test of 'G' by H. S. Perera. Published from the Government Training College, Colombo, 1938 (in English and Sinhalese). This is the most original of the tests so far worked out for India. This test is about midway between verbal and non-verbal. (For middle and high schools.)

An Urdu Adaptation of Ballard's Group Test for Juniors by Jalal Udin. See article entitled 'An experiment in Intelligence Testing' in *The Punjab Educational Journal*, Vol. XXXIII, No. 10, 1938. (For middle schools.)

Hindustani Group Tests of Intelligence by S. M. Mohsin. See article 'Intelligence and Its Measurements' in *The Indian Journal of Education*, Vol. II, Nos. 9-11, 1938. (For use in middle schools.)

Group Tests of Intelligence for Bengali Scholars (Class V-X) by H. C. Banerjee, D. N. Roy, and T. P. Bhowmik, Dacca.

A Composite Test along the lines of the *Ballard Test* under the auspices of the David Hare Training College, Calcutta.

The Ballard Test in Gujarati under the auspices of the Karachi Training College.

Composite Tests in English, Tamil and Telugu under the auspices of the Lady Willingdon Training College, Madras.

The Mahalanobis Composite Tests in Bengali under the auspices of Calcutta University.

Maitl's Composite Tests Forms A and B, in Bengali and Hindi under the auspices of Calcutta University.

An Adaptation of the Terman Intelligence Tests in Bengali by the Psychology Department of Calcutta University

An English adaptation of one of Cyril Burt's verbal intelligence tests by Prof. D'Sylva of Spence Training College, Jubbulpore. This is used as an aid for selection of college students where ability to understand English is a most important requirement. When used in this way by students whose mother-tongue is not English, it must be borne in mind that the test is not a true intelligence test but a combination of intelligence and ability to understand English test. English norms should not be used nor comparison with students whose mother-tongue is English made without due allowance. There are still others in preparation

NON-VERBAL GROUP TESTS

Very different from any of the tests hitherto described is the Goodenough 'draw-a-man' test. The child is instructed to make the best possible drawing of a man. Artistic ability is not scored but such points as indication (no matter how crudely) of feet, ears, eyes, nose, mouth, right number of fingers, inclusion of clothing, hair, etc. This test has in America a correlation of between 0.55 and

0.78 with the Stanford-Binet test which means that it is a test of real worth. Very little adaptation needs to be done for its use in India as the only language used is seventy-five simple words in the direction to the examinees. The only materials required are paper and pencil.

This test was standardized for India by the present writer.¹ Norms are given for both age and school placement based on 2,600 examinations. All the examiner needs to conduct this test is a copy of the Indian norms and the book describing Dr Goodenough's original experiment. This includes the manual for scoring.²

The advantages of this test are: (i) freedom from the use of language; (ii) it is very quickly scored; (iii) it can be given to children from about five upward, if they know how to use a pencil; (iv) it can be given to large groups at once.

The disadvantages of this test are: (i) it takes considerable skill to evaluate the drawings correctly; (ii) the scoring is more subjective than that of most intelligence tests; (iii) although Dr Goodenough found that the kind of drawing instruction given in primary schools did not influence the score appreciably, the widespread lack of drawing practice in India, does seem to handicap many children disproportionately; (iv) if the pupil does not do his best or does not understand what is wanted, the test is entirely invalidated for that pupil. This handicap can easily be overcome if such pupils are re-examined later and given personal attention.

This test usually gives a fairly valid judgement but when it does miss it is likely to miss much more completely than longer tests. Therefore individual results should

¹ *Teaching*, June 1935. See examples on p. 174.

² Goodenough, F. L., *Measurement of Intelligence by Drawings*, George G. Harrap & Co., London, 1926.

never be used unsupported by other information. As a survey test of larger groups its validity is quite high.

The correlation of this test as used in India was 0.55 with Kamat's Binet Scale, which is fair. Its reliability is 0.92 by the test and re-test method.

This test is usable in America for children aged 5 to 12. In India it is usable up to the age of 15 except in the case of bright children.

EXPLANATIONS TO ILLUSTRATIONS

No. 1. By an 8 year-old in the second grade of school. IQ 100, or just average (on the Indian scale).

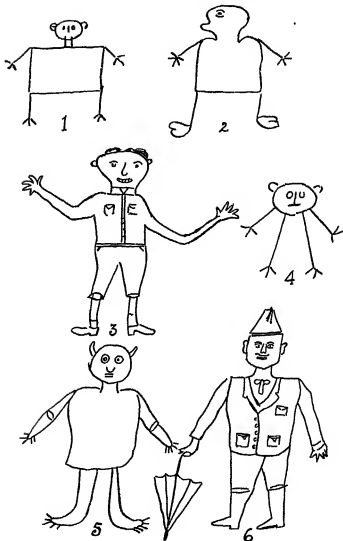
No. 2. This drawing is slightly more developed. IQ 112, second grade (on the Indian scale).

No. 3. What a contrast! There is understanding of the relationship of the body parts and clothing. The arms and legs have dimension, the number of fingers is correct. On the American scale this boy has an IQ 138 and on the Indian 200. He is in the third grade and though only 8 is the best pupil in the class.

No. 4. This and the following drawings are by 11-year-olds. This drawing is more primitive than those by the 8-year-olds. No body is shown, the arms stick out of the head, the number of fingers is incorrect. IQ 54 (Indian scale). This boy has attended school for nearly six years and is still in the second grade. It is clear that he cannot profit from school attendance.

No. 5. IQ 110, but notice that the 11-year old's drawing is more developed than the 8-year-old's (No. 2) with the same IQ.

No. 6. A very superior drawing for an eleven-year-old. No real drawing ability is shown, but notice the correct detail and imaginativeness. Nearly all the possible points on the scale have been scored. IQ 125 on the American scale and about 200 on the Indian.



Very recently an Intelligence test was developed in East Africa by Dr R. A. C. Oliver,¹ research worker of the Carnegie Foundation, which compares with verbal tests in effectiveness but has the advantage of being nearly independent of reading and writing skill. This has been adapted for use in Hindi by the author of this book.²

The only reading and writing skill required is that of reading and writing the letters of the alphabet and the figures 0 to 9. Thus pupils coming from the most backward schools or with the most meagre education can be tested. The preliminary practice exercises given to examinees to insure their understanding of what is to be done, are exceptionally well worked out and are designed especially for use in areas where such tests are unfamiliar. The test is probably as fool-proof as one can expect any Intelligence test to be. It is strictly objective. Only the speech of the examiner needs to be translated although test sheets have to be reprinted for each area having a separate written alphabet. With such a slight change comparisons between the various language-areas will be easily possible. In Anglo-Vernacular schools the use of the English alphabet makes the same test sheets usable all over India although the use of the test should not be restricted to Anglo-Vernacular schools.

Norms are available for both age and school grade. The reliability is 0.88.

The five parts of this test are picture numbering, similarities and differences, picture completion, number series, and picture absurdities.

Of all the comprehensive tests that have been tried

¹ *General Intelligence Test for Africans* (see p. 147). See also *The Testing of Intelligence*, pp 165-75.

² See *Teaching*, X, 3 and 4, March and June 1938. The plan and directions of this test are reproduced in Appendix C.

out in India thus far, this test is undoubtedly the most easily adaptable in all language-areas. The writer can place the blocks for the illustrations at the disposal of responsible parties who wish to have it reprinted for use in their language area. If this test lives up to expectations in being a valid measure of intelligence, it can soon become an inter-lingual test for use all over India.¹

In order to make this test more effective in Classes VII and VIII, a verbal supplement of four parts (sentence tangles, analogies, synonyms and opposites, arithmetic reasoning) has been added.

The use of the verbal supplement is optional. It is much more dependent on schooling than the non-verbal part. This is available in Hindi only, and will not prove to be easily translatable or comparable to tests in other languages.

As the second edition of this book goes to press the *Union Christian Training College Non-Verbal Group Test* (Berhampore, Bengal) has been got ready for standardization. It was prepared by T. C. Vicary and M. Draper. This test enjoys as thorough a preliminary experimentation as any group test in India. It is intended for ages 8 to possibly 15. It has succeeded remarkably well in eliminating all language and script factors so that with but a few strokes of a pen copies can be adapted for use in any language area. This is a very great advantage which is enjoyed by no other group test in India to the same extent. Both the *Perera Test* and the *General Intelligence Test* are easily adapted to other language areas without

¹ This test has now been adapted into Urdu by R. Ewing and G. Leeder. Copies may be obtained from Mr Ewing, A. P. Mission High School, Dehra Dun. It has also been adapted into Gujarati by Miss E. Chilson. Copies are obtainable from The Superintendent, Methodist Mission Training School, Godhra, Panchmahals.

change of standard but require reprinting when used in a different language area. The *U.C.T.C. Test* is free from this handicap.

Probably the most consistent use in India of Intelligence tests has been made by Mr J. H. Warnshuis and Dr Mason Olcott of Vellore.¹ Since 1922 Dr Olcott has been making up a new battery of tests every year. Of course the battery is not standardized with norms but the method suits admirably the use for which it is intended. Dr Olcott is entrusted with selecting 50 out of about 250 applicants each year for middle and high school education. He first gives tests in arithmetic and silent reading of the Standardized Test style. Too great credence is not placed on these two tests, however, as the pupils come from a variety of primary schools where teaching and general advancement are far from uniform. Then a Group (mostly non-verbal) Intelligence test of 6-9 parts is given. All the tests used are changed every year in order to eliminate coaching effects. During the course of years a great many types of tests have been utilized. Undoubtedly Dr Olcott has much valuable experience on what types of tests are most satisfactory for use in India for children aged 10-14. Anyone contemplating making up group tests would do well to consult him.

This experiment is worth serious attention on the part of those who have to plan scholarship examinations. This method is nearly as superior to the old-fashioned scholarship examinations as the motor-bus is to the ox-cart. Norms are not needed in such a testing programme as the 50 out of 250 with the best scores are chosen. Coaching is better guarded against than under the old-fashioned

¹ Warnshuis, J. H., 'The Arcot Mission Experiment in Tests'. Reprinted from *Christian Education* (January 1924). Obtainable from Dr Olcott, Vellore, S. India

examination and two entirely separate indexes of the child's ability are before the examiner, namely success in the two most important school subjects and natural ability not dependent on the type of teaching hitherto received. This type of examination takes less time than the government scholarship examinations and is infinitely more informative. This experiment has had the helpful criticism of such eminent authorities as Thorndike and Pintner.

FURTHER NEEDS

There is room for several reliable and valid tests for each age-group. Wherever the Binet scale has been adapted a most trustworthy measure of intelligence from age 3 to adulthood is available.

In addition to an individual test, group tests for all ages are needed. It seems as if suitable tests for late high school and colleges are either under way or will be forthcoming along the lines of the verbal tests described. Perhaps one may be devised which uses English as a medium but demands such a limited use of English that the college student is not handicapped greatly because English is not his mother-tongue. Such a test could be used all over India instead of being confined to one language-area or undergoing such changes in translation as to make the test of the different language-areas hardly comparable. But the English used must be much simpler than found in imported tests.

For middle and lower high schools there are now verbal group tests for several language-areas as well as three non-verbal group tests which can easily be adapted to different language-areas, one with little or any trouble (See page 176.)

A good group test for children under 8 does not seem

test at present which deals with very young children but this is not sufficient in itself. A more general test is needed. This will be more difficult to work out than any of those mentioned above and will take considerable experiment. It will have to be non-verbal and very general in scope.

The author has experimented with a short individual test for pre-school and Class I pupils but is not, at the time of writing, too hopeful of the test being valid enough to be useful.

The test employs the usual devices used in Kindergarten tests in America and England such as pictures, comparisons, picture completion, picture absurdities, etc. However, Indian children in villages are too little acquainted with pictures to do justice to themselves in such a test. It was found that half a year of school influences the test score about as much as two years of development outside the school. In other words, school influences the score which makes it invalid for comparing children whose school or cultural experience differs.

The matter of a test for ages 5-8 is important for compulsory education has already accentuated the problem of the retarded child. Formerly the retarded child was simply failed and failed until he left school. Under compulsory education teachers are faced with the problem of what to do with the child who simply cannot keep up with his class but is compelled to keep on attending. Attendance officers will have to have a means of knowing which children should be excused from the compulsory act. For this a simple and nearly fool-proof Intelligence test is needed. Whether such a test will be devised that can safely be left to the discretion of the rank-and-file attendance officers, even after they have received some training, remains to be seen.

VIII

PROGNOSIS AND APTITUDE TESTS

My dictionary defines 'prognosis' as 'any prediction or forecast; foreknowledge'. A prognosis test is a test which attempts to foretell the amount of a person's probable success in some certain line. A prognosis test in geometry, for instance, seeks to tell a student whether he will do well in this subject and be rewarded by getting something out of it, or, that he will do poorly in this subject and be well advised to leave it alone and take something he is better fitted to take. Another prognosis test might tell the student what offers him better prospect of success.

Such tests would undoubtedly be of great value not only in school but in the choice of vocation and in many other situations in life. They would help prevent many disappointments and even tragic failures and would help all of us to make better use of the many or few talents we have. But the argument is not whether or not such tests would be useful but whether there are tests that can predict probable success or failure with any accuracy.

Every valid Intelligence test is a prognosis test in a general way. We know that the higher your IQ is the better are your prospects of passing geometry, language, mathematics and other tests; also the better your prospects are of being well liked and generally successful in life. But we also know that highly intelligent people may be very poor in some things. Some Ph.D.s should never be permitted to take a wrench in their hands when near a motor car for fear of the damage they will do. The trouble

is not their lack of knowledge about wrenches or motor cars but their lack of mechanical common sense, which they do not realize. Some very bright people should not be permitted to drive a motor car because no matter how much they practice driving they lack ability to judge accurately the speed and distance of moving vehicles. A general Intelligence test has some predictive value as to our probable success in most things but it is worthless for predicting success in motor driving, piano tuning, beautiful handwriting, dancing, or speed in typewriting. For specialized skills a special type of test is needed.

One of my friends wasted much money and much of his son's happiness in trying to make a violinist out of him. He would have saved trouble and expense had he had the *Seashore Measure of Musical Talent Tests*¹ given to his son before making up his mind that the boy should become a noted violinist, for later it was found out that the boy had a defective sense of rhythm and a defective ear which made keeping in time and in tune an impossibility. Such matters the *Seashore Tests* can tell before music lessons are begun. Another friend of mine studied medicine, almost completing his course before he discovered that the muscular control of his hands was so poor that he would never be able to do the simplest surgery. This restricted him to a narrow field in medicine which made him regret exceedingly that he had not taken up law for which he felt an equal attraction. His failing could easily have been told beforehand by a competent psychologist.

One of the first things to be done in an effort to foretell future success or failure is to locate physical or motor defects which make success in the chosen field impossible

¹ Published by the C. A. Gregory Co. The material for this test includes six gramophone records.

or difficult. It is also of value to list the physical assets. A bacteriologist's success is so dependent on good eyesight that it is certainly of value for anyone with ambitions to become a bacteriologist to take stock of his assets in this direction.

There are some subtle physical assets and liabilities which affect future success in numerous fields which the physician does not ordinarily discover. In an effort to discover and measure these the psychologist has composed so-called Aptitude tests. An Aptitude test is interested in the future just as much as a prognosis test, but functions in a more physiological field. The value of an Aptitude test, just as much as prognosis, is in whether it actually does indicate what the probabilities of future success are. The distinction between the two kinds of tests is not a hard and fast one.

Aptitude is quite different from ability. I probably have the proper aptitude to learn dancing but I have not the ability to dance. There is probably no reason why I could not have learned to dance but I haven't the ability to dance because I have not seriously tried to learn. Ability demands practice but the ability to do something cannot be developed unless one has an aptitude for it. A lame awkward person lacking a sense of rhythm has little aptitude for dancing. Aptitude comes from inborn capacity and physical condition whereas ability is the inborn capacity plus the effects of training (or practice). We all have the aptitude for doing a great many things that we have not taken the time to develop our ability to do. Thus the task of an Aptitude test is to find out the amount of inborn capacity there is for mastering certain skills regardless of the fact that among the people tested there will be some who have practised this skill far more than others. In Aptitude tests tasks have to be included which

the examinee is unlikely to have practised, even though he may have been practising the kind of work for which we wish to measure the aptitude. Such tasks force the examinee to exercise abilities for the first time. His performance in these tasks gives a clue to the amount of his latent aptitude.

Mechanical ability has a correlation of only about 0.20 with general intelligence. Thus many children who do not get along well in school may still make tolerable mechanics. It is not as necessary to find these in India as it is in a highly industrialized country, but still, since a Mechanical Aptitude test, if it is a valid one, will do for the prospective semi-skilled mechanic what a general Intelligence test does for a prospective student of a high school, it would be convenient to have such tests in India. With the demand for bifurcated education there will be a demand for such tests in India. There has been some attempt on the part of psychologists in India to measure mechanical aptitude in individual cases and the Department of Education in at least one province (probably more) has appointed a committee to work out tests which can be used in vocational guidance on a larger scale. But it does not seem as though any positive recommendations have as yet been submitted for the use of educationists.

One of the earliest tests along this line is the *Stenquist Assembly Test of Mechanical Ability*¹. This consists of giving the examinee a series of ten mechanical contrivances which he must put together rapidly in order to get a good score. Such a test assumes some knowledge of mechanics such as a boy in an industrial civilization has opportunity of acquiring for himself if he is mechanically

¹ World Book Company.

bent. Such a test could not be used in the case of village boys in India who have had little opportunity of seeing and handling mechanical devices.

*The MacQuarrie Test for Mechanical Ability*¹ puts no premium on knowledge. It measures mostly speed of muscular reaction, skill of the hand and co-ordination of hand and eye. The first test in this battery gives a large number of parallel lines in each of which there is a very small gap at height levels which vary in each line. The examinee is asked to draw a continuous snake-like line passing through all these gaps without touching the edges of the gaps. As there are 84 such narrow gaps and a very short time is allowed, it takes co-ordination of hand and eye and excellent control of the muscles of the hand to achieve a good score. In the second test the examinee is required to put a series of three dots inside each of a series of small circles at the greatest possible speed. Another test concerns itself with copying simple angular designs within an ordered series of dots. These dots are of some help to the examinee but are of even more help to the examiner because a mistake in visual judgement is quickly revealed. Still another test in this battery is called 'pursuit'. A series of curved, thread-like lines (like a badly tangled fishing-line cut through the middle of the tangle) are placed before the examinee. Some of the 'cut' ends have numbers and each thread has to be followed through and the correct number put against the opposite end. About forty threads must be unravelled in three minutes. There are altogether eight tests within the battery. Such a test as this could have given a hint of future trouble to the young man who studied medicine only to learn that his hand could never develop the required skill for surgery.

¹ C. A. Gregory Co.

Such a test as this is of aid in identifying those of slow and uncertain hand and eye, and is extensively used in examining prospective typists and stenographers and those who need to use hand and eye in rapid and fine work

The National Institute of Industrial Psychology in London has published a test for Clerical Ability which is acknowledged to be highly reliable and valid. This Institute is conducting elaborate experiments in measuring mechanical aptitude and has a number of noteworthy successes to its credit

Just at present the experts are not agreed as to what the prospects for Aptitude tests are. Some psychologists have hoped that tests can be devised that will show whether a certain boy would do better as blacksmith, carpenter, or tinsmith. But there is little evidence that such fine distinctions in aptitude can be made. Psychologists are however prepared to say that certain boys should not try blacksmithy, carpentry, or tinsmithy and that certain others should not try watchmaking, typing, or dentistry. This already is quite helpful. At present industrial psychology is still in its infancy and scientific evidence on the value of its recommendations still needs to be accumulated.¹

This is not a field for the novice. Vocational guidance is urgently needed in India but those who enter the field will have to be carefully trained. There is no material as yet available that is fool-proof enough to put into the hands of any but the expert.

¹ Cattell claims that the best mechanical aptitude test available at present is the recent one by Dr J W. Cox of the National Institute of Industrial Psychology. He also speaks highly of the *Minnesota Mechanical Aptitude Test* (Educational Test Bureau, University of Minnesota)

We may now return to the subject of prognosis tests. A discussion of the fact that mechanical ability is not to any considerable extent a matter of general intelligence should have prepared us for consideration of the theory that there are some limited abilities which are quite different in quality from general ability. These may be developed in us either more or less easily than the general run of our abilities indicates. This is the 'G' and the 'S' of which Spearman speaks (see p. 145). There are some people of high intelligence who simply cannot understand mathematics. There is a story in circulation that the great mathematician, Einstein, is very weak in simple arithmetic despite his most eminent success in higher mathematics. Whether this is true or not, psychology would by no means be surprised at a mathematical genius who is poor in arithmetic. Higher mathematics and arithmetic require different abilities. Darwin was considered a dull student in all subjects but science and in this subject he was considered more painstaking than brilliant. Yet he displayed great brilliance later in life. Language studies and science take different types of application on the part of pupils, as also do geometry and physics. Aptness in one subject usually means aptness in others but the exceptions to the rule are so many that it is well not to be bound by the rule.

Just how can one find out who will do well in geometry or Sanskrit before forcing students to spend a year or two trying to master the subject? The technique used in trying to find the answer to this question is as follows.

The experimenter chooses several hundred students just about to begin the study of geometry. He gives them several Intelligence tests choosing those which require a wide variety of performances. He will at the same time test the pupils thoroughly in the kind of abilities which

seem to be closely related to the study of geometry and in which the pupil already has some experience, such as recognition of geometrical forms, computation of areas and volumes, and ability to do simple constructions. The experimenter will then wait an entire year, or until the study of geometry has been completed, and will then give a very comprehensive test in geometry. With the knowledge of the pupil's success in studying geometry before him he can then compare this success with the pupil's success in the various types of tests given the year before. This comparison is made by means of the correlation of the results in geometry and the various other tests. How this is done is briefly described in Chapter X (p. 206). The experimenter can thus see how later success in geometry compares with skill in computing areas and volumes, with ability to do word analogies, completions, etc., without accepting anyone's opinion as to what abilities are related to a capacity for learning geometry. With this knowledge of the relationship between successes in other subjects and later success in geometry he is then prepared to list the things which correlate highest with success in studying geometry. He is justified in assuming tentatively that the same condition will continue. Thus armed he is ready to choose items and prepare his success-in-geometry-prediction test.

When this is ready there remains to be made a careful try-out of his prognosis test. If it works, it works; and if it doesn't, it doesn't. Whether it does or not can only be told after widespread use and careful checking on how the predictions turned out.

During the experimental stage of the test, no pupil should be debarred from studying geometry because he did not do well in the prognosis test. The test is only an experiment and may not do what it is expected to do. If

after widespread experiment it is proved that the test can identify the student who will succeed or fail better than a general Intelligence test can, it is valid enough to be used by student advisers.

There are a number of prognosis tests already in the field; in geometry, modern languages, ancient languages, reading, etc. At present prognosis tests must be considered experimental. It is still too early for a considered opinion as to the success of any one test. Symonds claims that although any good general intelligence test is a good prognosis test, success in making tests that do better for special subjects than the general intelligence tests, is a movement which, although it gives considerable promise is still in its infancy.

A prognosis test in teaching ability is used in America in some teachers' training schools to aid in selecting the most apt pupils from the large number of applicants for admission to the school. This test has been proved to be helpful in choosing those who have the greatest likelihood of success in the studies constituting the course given in the training school, but a further trial for the test remains. Careful note will have to be taken of the success in actual teaching of each examinee for some years after his graduation and entry into the teaching profession. Not until these results are available will the real value of the test become known.

There are, at the time of writing, no prognosis tests developed for use in India. There is however a *Professional Test for Teachers* described by its author, Dr J. Manry, as 'an occupational fitness' test.¹ This is neither

¹ Published by the Bureau of Educational Research, Allahabad Christian College. The results with this test in the Lahore Training College were described by Mr Pars Ram in *The Indian Journal of Psychology* (1937).

an aptitude nor a prognosis test but since there is no separate chapter on occupational tests and this test certainly merits mention, it is included here.

This test concerns itself with only one aspect of occupational fitness, namely, intellectual background. Knowledge does not necessarily make a good teacher but lack of knowledge certainly disqualifies for a post in which a competent and well-informed teacher is desired. The test is composed of the following parts:

PART I. General Knowledge (acquaintance with commonly used articles, phrases etc.).

PART II. General Knowledge (familiarity with the names and achievements of prominent contemporary personages).

PART III. Professional technique

PART IV. Professional judgement.

PART V. Educational leaders (Indian and foreign).

PART VI. Educational literature (Indian and foreign).

PART VII. Educational terminology.

This test is biased in favour of the well-trained teacher of up-to-date outlook. An occupational fitness test of this sort may be expected to be biased, for its purpose is to select candidates of a certain type. The progressive manager or principal of a school is not interested merely in finding someone who can teach according to any method, no matter how archaic or inefficient it may be, but who will fit in with the viewpoint and aims of the institution in which his services are required. The test frankly aims at locating the teacher with a certain professional background and outlook which is considered necessary for efficient teaching of the type desired.

Similar to the above is the *Teaching Aptitude Test* by Dr L. K. Shah of Lucknow Christian College. A name like *Professional Test for Teachers* would be more apt. Its five parts deal with general information, general

intelligence, educational practice, general information and professional judgement. This test, too, should prove a help in the preliminary appraisal of applicants for teaching posts and as a test for students in training schools. Both tests are in English.

IX

MEASURING OTHER QUALITIES

It is only natural that the first attempts to measure scientifically the abilities and aptitudes of human beings should have concerned themselves with the narrowly mechanical abilities. It is much easier to measure memory span than to measure the ability to appreciate the good qualities of a rival. It is easier to measure objectively keenness of eye or ear than to measure keenness of moral perception. Psychology borrowed the methods and standards of measuring sensory response, reaction time, sensitiveness to pain, etc., from the older sciences of physiology and physics; but when it wanted to measure imagination, will, sense of humour, prejudice, sociability, etc., it had to develop its own technique in a field more complicated than that in which physics and physiology function.

We find Binet dissatisfied with the narrow field in which psychology was then attempting its measurements and asking: 'How can we measure richness of inspiration, accuracy of judgement and general ability of the mind?' He attempted to answer the question and started the world well on the way to finding the answer. It is only forty years since Binet put that question, a comparatively short time in the life of a science, and already substantial progress has been made in measuring fairly complicated and subtle human qualities. To be sure, psychology cannot measure as accurately as physics but the human mind is not the simple process that gravity or electricity or air pressure are.

Many measurements in the psychological field need to be performed in the psychologist's laboratory. Our concern in this book is only with such measurements as have been put into test or rating-scale form and are at the disposal of others besides the professional. A few titles selected at random from publishers' catalogues are listed simply to give an idea of the varied field in which they operate.¹

Character Inventory Chart by B. L. Dougherty, F. K. O'Reilly and M. E. Mannix. Public School Publishing Co.

'Telling what I Do' Tests by H. E. Baker. (A test of behaviour attitudes). Public School Publishing Co.

Sims Score Card for Socio-Economic Status by V. M. Sims. Public School Publishing Co.

Torgerson Diagnostic Teacher Rating Scale of Instructional Activities by T. S. Torgerson. Public School Publishing Co.

Dragoo Rating Scale for Shop Teachers by A. W. Dragoo. Public School Publishing Co.

Downey Individual Will Temperament Tests by J. E. Downey. World Book Co.

Downey Group Will Temperament Test. World Book Co.

New York Rating Scale for School Habits by E. L. Cornell. World Book Co.

Haggerty-Olson-Wickman Behavior Rating Scales by M. E. Haggerty, W. C. Olson and E. K. Wickman. World Book Co.

Sense of Humor Test by J. C. Almack. C. A. Gregory Co.

Tests of Appreciation of Poetry by A. Abbott and M. R. Trabue. C. A. Gregory Co.

Tests of Appreciation of Literature by A. Abbott and M. R. Trabue. C. A. Gregory Co.

A-S Reaction Study in Personality by G. W. Allport. Houghton Mifflin.

¹ The reader is again referred to Cattell's *A Guide to Mental Testing* for introduction to more recent experiments and an appraisal of their value.

- Allport-Vernon Study of Values* by G. W. Allport and P. E. Vernon. Houghton Mifflin.
- Harper Test in Social Beliefs and Attitudes* by W. H. Harper. Public School Publishing Co.
- Morris Trait Index* by E. L. Morris. Public School Publishing Co.
- The Pressey X-O Tests for Investigating Emotions* by S. L. Pressey. C. A. Stoelting Co.
- Tests of International Attitudes* by R. Neuman. Association Press, New York.
- Watson's Tests for Fairmindedness* by G. B. Watson. Bureau of Publications, Teachers' College, New York.
- Freyd's Occupational Interest Test* by M. Freyd. Bureau of Publications, Teachers' College, New York.
- Strong's Vocational Interest Test* by E. K. Strong. Bureau of Publications, Teachers' College, New York.
- Bernreuter's Personality Inventory* by R. G. Bernreuter. Stanford University Press.
- Thurstone Neurotic Inventory* by L. L. Thurstone. Stanford University Press.
- Bernreuter's Self-Sufficiency Test* by R. G. Bernreuter. Stanford University Press.
- Laird Introversion Test* by D. A. Laird. Colgate University Book Shop, New York.
- Allport Ascendence-Submission Reaction Studies* by G. W. Allport. Houghton Mifflin.

It will be noticed that many of the above titles are rating-scales, inventories, or questionnaires rather than tests. For a time rating-scales and questionnaires were rather discredited but with an improved way of using these devices and a better understanding of how not to use them they are receiving widespread recognition.

Performance tests, that is tests which judge according to how the examinee performs in certain standardized situations, are rather difficult to use in measuring conduct and personality. Hartshorne, May and Shuttleworth collaborated in making one of the most elaborate measure-

ments ever made in any psychological field ¹ In order to test honesty and deceit they actually arranged to have the teacher send pupils to a shop to purchase some article. According to the arrangement with the shopkeeper the pupil was given more change than he should have been given. In rendering account of the transaction to the teacher, there was an opportunity for dishonesty which could easily be scored. Numerous other opportunities were given for cheating in examinations, cheating in games, lying, and taking unfair advantage. Through these tests it was found that most of the examinees were honest in some things but dishonest in others; that honesty is not a single virtue which is applied in all situations but a series of independent habits specific to various situations. Such elaborate arrangements and such a great expense was involved in carrying out these tests that the methods used by Hartshorne, May and Shuttleworth, though very valuable in research work, are not generally practical for testing purposes. Performance tests show little prospect of being generally useful in measuring character because of the difficulty of administration.

The rating-scale, seemingly a very subjective way of measurement, has been refined enough to produce fair objectivity. Supposing we wish to find out the personality traits of a certain person, the judgement of several people is worth more than the judgement of one. When we have the judgements of ten people regarding this person we are in a position to generalize and come to conclusions which will be more reliable than the opinion of an individual. We can aid those who are about to make the judgement by outlining a detailed plan of rating which gives the utmost help in making a subjective estimate.

¹ *Studies in Deceit*, 1928, and *Studies in the Organization of Character*, 1933. The Macmillan Co., New York.

as objective as possible, as is evident in the example to follow.¹ A clever graphic rating-scale usually results in fair agreement among the raters.

GRAPHIC RATING-SCALE

ALERTNESS

Consider whether (s)he is inclined to be absent-minded, or whether (s)he is always on the spot

Is very wideawake and keen.	Shows ready attention.	Responds well on the whole.	Not very quick	Dull.
-----------------------------	------------------------	-----------------------------	----------------	-------

PERSEVERANCE

Consider power of overcoming difficulties

Forceful; never dismayed however insuperable the obstruction.	Makes considerable attempt to overcome difficulties	Normal	Needs some encouragement from others.	Easily gives in in spite of encouragement.
---	---	--------	---------------------------------------	--

ATTENTION

Consider concentration of attention

Abstracted; difficult to obtain attention to tests	Distracted by own ideas or by chance extraneous stimuli, but attention readily reverts to tests.	Normal attention to other stimuli may be shown, but not such as to impair efficiency in doing the tests.	Attention to test undisturbed by chance extraneous stimuli.	So pre-occupied with tests that extraneous stimuli are definitely irritating.
--	--	--	---	---

¹ Hunt, E. P. A. and Smith, P., *Teachers' Guide to Intelligence and Other Psychological Testing*, p. 73. Evans Bros., London.

COMPREHENSION

Consider ease with which test instructions are grasped

Compre- hends only after un- usual elab- oration of instruc- tions.	Grasps ideas with repetition of instruc- tions.	Required no repeti- tion of in- structions.	Grasps ideas be- fore usual instruc- tions are completed.	Seems to compre- hend on production of test material.
---	---	--	--	--

There are also self-rating scales or inventories to be had. A person's self-estimate on a self-rating scale may be compared to the rating others give the same person. The results are highly interesting and to the psychologist very revealing. Self-rating scales may also be used unsupported by other ratings and are made much use of in student guidance and personnel work. When unsupported by other ratings—they must however not be regarded as measurements, but as subjective self-estimates that aid a person in looking at himself somewhat objectively, particularly if these self-estimates are discussed with an adviser—they are nevertheless too subjective to be considered an objective measurement.

Such an inventory is the *Dougherty, O'Reilly-Mannix Character Inventory*.¹ This gives ten headings under each of which ten possible ratings are given. The pupil is asked to indicate the rating he considers most characteristic of himself. The headings are health, honesty, loyalty, cheerfulness, courtesy, co-operation, moral courage, industry, self-control, and leadership. The first mentioned quality in each column has a value of ten, the second mentioned quality a value of nine, the third a value of eight, etc.

¹ Public School Publishing Co.

HEALTH	HONESTY	MORAL COURAGE	INDUSTRY
vigorous	honourable	confident	original
robust	fair	courageous	ambitious
heartty	true	decisive	dutiful
strong	frank	determined	thorough
well	honest	hopeful	steady
weak	unfair	timid	slow
delicate	untrue	doubtful	careless
ailing	deceitful	discouraged	neglectful
sickly	faithless	cowardly	lazy
ill	dishonest	hopeless	idle

An entirely different technique is used by Pressey¹ in his test for investigating the emotions. According to the author the purpose of this test is:—

1. To furnish aid in the study of delinquents, neurotics or other typical individuals.
2. To furnish aid in the study of sex and individual differences, affectivity and moral judgement, and emotional make-up.

PART I

Directions

Read through the twenty-five lists of words given just below and cross out EVERYTHING THAT YOU THINK is WRONG—everything that you think a person is to be blamed for. You may cross out as many or as few words as you like; in some lists you may not wish to cross out any words. Just be sure that you cross out everything you think is wrong.

1. Begging, smoking, flirting, spitting, giggling.
2. Fear, anger, suspicion, laziness, contempt.
3. Dullness, weakness, ignorance, meekness, stinginess.
4. Fussiness, recklessness, silliness, nagging, fibbing.
5. Extravagance, sportiness, boasting, deformity, talking-back.

¹ *The Pressey X-O Tests for Investigating Emotions.*

PART II

Directions

Read through the twenty-five lists below and cross out EVERYTHING ABOUT WHICH YOU HAVE EVER WORRIED, OR FELT NERVOUS OR ANXIOUS.

1. Loneliness, work, forgetfulness, school, blues
2. Sin, headache, fault-finding, sneer, depression
3. Meanness, clothes, sickness, looks, unfairness.
4. Temper, disease, pain, money, awkwardness.

PART III

Directions

Read through the twenty-five lists just below and cross out EVERYTHING YOU LIKE OR ARE INTERESTED IN. You may cross out as many or as few words as you wish; there may be some lines in which you will not wish to cross out anything. But be sure you cross out everything you like.

1. Fortune-telling, boating, beaches, mountains, vaudeville.
2. Camping, tennis, hiking, eating, amusement-parks.
3. Beethoven, Edison, Napoleon, Raphael, Tennyson.
4. Kissing, flirting, pretty girls, talkative girls, athletic girls
5. Studying, dancing, day-dreaming, walking, reading.

There are twenty-five such lines under each of three parts.

The layman will probably be confused as to what such a test means to do. It is not a test in the sense in which the word is used in school parlance. In such tests there is no passing or failing mark. Such tests are for psychological information about the individual and not to measure prowess. It takes a fair understanding of psychology to make use of tests of this type.

On a similar subject but of the so-called 'adjustment questionnaire type' is the *Woodworth Psychoneurotic*

*Inventory.*¹ The questions in this questionnaire are built up around the many symptoms of the various types of neuroses. There are 116 questions in the inventory from which the following items have been chosen to indicate the general tenor. It goes without saying that it takes a psychologist to interpret the results.

- | | | | |
|-----|--|-----|----|
| 1. | Do you usually feel well and strong? | yes | no |
| 2. | Do you usually sleep well? | yes | no |
| 3. | Are you frightened in the middle of the night? | yes | no |
| 4. | Are you troubled with dreams about your work? | yes | no |
| 6. | Do you have too many sexual dreams? | yes | no |
| 8. | Do you have the sensation of falling when going to sleep? | yes | no |
| 10. | Do ideas run through your head so that you cannot sleep? | yes | no |
| 14. | Do you often have the feeling of suffocating? | yes | no |
| 16. | Are you bothered much by blushing? | yes | no |
| 30. | Did you have a happy childhood? | yes | no |
| 31. | Were you happy when 14 to 18 years old? | yes | no |
| 32. | Were you considered a bad boy? | yes | no |
| 33. | As a child did you like to play alone better than to play with other children? | yes | no |
| 34. | Did the other children let you play with them? | yes | no |
| 35. | Were you shy with other boys? | yes | no |
| 37. | Did you ever have a strong desire to run away from home? | yes | no |
| 38. | Has your family always treated you right? | yes | no |
| 42. | Do people find fault more with you than you deserve? | yes | no |
| 43. | Do you make friends easily? | yes | no |
| 44. | Did you ever make love to a girl? | yes | no |
| 45. | Do you get used to new places quickly? | yes | no |
| 54. | Have you ever seen a vision? | yes | no |

¹ Published by C. H. Stocking Co

- | | | | |
|-----|--|-----|----|
| 58. | Are you ever bothered by the feeling that people are reading your thoughts? | yes | no |
| 61. | Are you troubled with the idea that people are watching you on the street? | yes | no |
| 62. | Are you troubled with the fear of being crushed in a crowd? | yes | no |
| 63. | Does it make you uneasy to cross a bridge over a river? | yes | no |
| 68. | Do you worry too much about little things? | yes | no |
| 69. | Do you think you worry too much when you have an unfinished job on your hands? | yes | no |
| 79. | Do you feel like jumping off when you are on high places? | yes | no |
| 89. | Are you troubled with shyness? | yes | no |

A criticism commonly levelled at the questionnaire as a method of measurement is that people may fill out such a questionnaire in a spirit of fun and levity purposely falsifying answers and thus invalidating results. This criticism has been investigated and very few cases were found where the questionnaire was given in good faith and was not filled in in good faith. Of course the occasions on which such questionnaires are given must be carefully chosen.

Attitude questionnaires have of late become very popular in America and show promise of being quite useful in finding out the attitude of the public on all sorts of public and private questions, and are of value to the student of sociology, economics, politics, and the like. It is quite easy, however, for the questioner to come to false conclusions if the individuals who take part in the experiment are not the representative group the examiner takes them to be.

Questionnaires need not be in question form. A ques-

tionnaire can be of the True-False, Multiple Choice, or Classification style, just as well as of the question style.

The following example of an attitude questionnaire is a few items selected from L. L. Thurstone's *Scale of Attitude toward the Movies*.¹

The scale-value of each statement is shown in parentheses following its serial number. The higher the scale-value the more favourable the statement toward the movies.² The statements have been arranged in random order.

1. (1.5) The movies occupy time that should be spent in more wholesome recreation.
2. (1.3) I am tired of the movies; I have seen too many poor ones.
3. (4.5) The movies are the best civilizing device ever developed
8. (2.7) Movies are just a harmless pastime
9. (1.4) The movies to me are just a way to kill time.
10. (4.0) The influence of the movies is decidedly for good
12. (3.9) Movies increase one's appreciation of beauty.
14. (2.4) Sometimes I feel that the movies are desirable and sometimes I doubt it.
17. (4.3) The movies are the most vital form of art today.
18. (3.6) A movie is the best entertainment that can be obtained cheaply.
21. (1.3) Going to the movies is a foolish way to spend your money.

To return again to the question of testing morals, which was discussed in connexion with performance tests, one can test *conduct*, *moral knowledge*, and *moral judgement*. Knowledge and judgement of right conduct can be no

¹ Symonds, *Diagnosing Personality and Conduct*, p. 229 The Century Co., New York, 1931.

² 'Movies' is the popular American term for the cinema

means be accepted as being the same as right conduct itself, yet there is a certain value in knowing whether the examinee can make proper moral discriminations. One cannot expect moral conduct in a large variety of situations unless there is a sense of proper moral values present. One may know these values and deliberately flout them, or even fiendishly set out to do the opposite of the moral, but if one is lacking in judgement to recognize what is generally considered the most desirable line of conduct, this will reveal itself in a well constructed test. In a well constructed psychological test the examinee cannot guess what answer the examiner wants consistently enough to deceive him. A few test items chosen from different tests of knowledge and attitude on moral questions are given to illustrate the method used in such examinations. In these cases they happen to be Multiple Choice tests. The response the pupil considers most desirable is to be underscored

1. If another pupil wants to copy your work and hand it in:
 - (a) Let him do it and say nothing about it
 - (b) Let him do it and tell the teacher
 - (c) Don't let him do it and say nothing
 - (d) Don't let him do it and tell the teacher he wanted to.
 - (e) Don't let him do it and tell him you disapprove of cheating.
2. You see broken glass in the street:
 - (a) Pick it up
 - (b) Do nothing about it
 - (c) Tell the policeman about it
 - (d) Try to find the one who did it.
3. If someone asks to borrow your pencil.
 - (a) Tell him it is broken.
 - (b) Tell him that you lost it.
 - (c) Tell him that you do not want to loan it.

4. If someone steals your lunch:
 - (a) Take someone's else's lunch and even it up.
 - (b) Report it to the teacher.
 - (c) Cry about it.
 - (d) Say nothing about it.
5. John noticed that everybody in the class was cheating on a test so he cheated too.
Say whether his act was right, wrong, or considering the circumstances, excusable.
6. Charles did not want to play marbles for keeps (the winner keeping the marbles) but the other boys laughed at him and called him a 'sissy' so he went ahead and played for keeps so they would not laugh at him. Was his act right, wrong, or excusable?

As mentioned in the beginning of this chapter, the higher mental and ethical qualities are harder to measure than mechanical ability. They are both harder to measure and the period since measurement has been attempted in these more complex fields is shorter. Relatively little that has been done in this field can be regarded as having passed through the experimental stage into the validated stage. In the hands of the careful psychologist there are a number of measures which are considered safe but there is little that can be entrusted into the hands of the untrained to use safely and accurately. This seems to be the consensus of opinion of various authorities. We quote Dr P. E. Vernon:

Temperament and character are too complex and fluid ever to be qualified by a few simple short cuts. But temperament testing may almost be said to have outgrown its childhood, to have discovered some of its own worst weaknesses, and to offer an abundance of ingenious techniques, many of which already show a moderate value.¹

¹ *Testing of Intelligence*, p. 130

Regarding character tests Symonds says.—

It is possible to measure knowledge and judgement with reference to conduct through the application of several useful tests which have been constructed for measuring health, Biblical knowledge, ethical knowledge, etc. These tests have very satisfactory reliability, comparing favorably with tests in the school subjects, similarly constructed. They correlate somewhat with each other and substantially with intelligence in general.¹

Again Symonds says:—

Performance tests have a real and valuable place at the present time in experimental work.. ... But there must be considerable further development before tests of this type become a feasible tool in clinical work.²

The reference from which the two last quotations were made is an excellent handbook on the type of measurements described in this chapter. It describes limitations as well as advantages of many more types of measurement than could be mentioned in this short chapter.

A very interesting and detailed description of some of the more significant experiments along the lines described are found in A. E. Wiggam's book, *Exploring Your Mind*.³ Wiggam is a popularizer of biology and psychology rather than a scientist himself and waxes more enthusiastic about some of the experiments than is probably justified, but his account will be found interesting, stimulating, and instructive.

¹ *Diagnosing Personality and Conduct*, p. 294

² *Op. cit.*, p. 354

³ Bobbs Merrill Co., Indianapolis, U.S.A., 1928.

SIMPLE STATISTICAL DEVICES AS A HELP IN INTERPRETING RESULTS

THE most confusing part of the modern testing movement to the novice is the new language the movement has acquired in its use of statistical devices. If it were not for the mathematical complications, New-Type testing and Intelligence testing would be more intelligible to the layman. And yet the mathematics are fairly simple (much of the mathematics of the high school course is more difficult). But statistical processes are unfamiliar even to many teachers of mathematics. Some of the statistical devices will be briefly described. For an adequate though simplified treatment the reader is referred to H. Rugg's *A Primer of Graphics and Statistics for Teachers*¹ or E. W. Tieg's and C. C. Crawford's *Statistics for Teachers*.² The purpose of this chapter is to give only a brief explanation of some of the more familiar terms and devices and to indicate with which it is important for the amateur to familiarize himself.

Much of the development of the testing movement and the development of research in psychology and education was only made possible by the recent refinement of the statistical treatment of the data. In the refinement of the statistical technique the English have made the most notable contributions.

Any teacher knows what 'rank order' is, even though

¹ Houghton Mifflin, New York, 1925.

² Houghton Mifflin, New York, 1930

he may not be familiar with the term. When a teacher arranges examination papers in order, putting the paper with the highest score on top, the paper with the second highest score second and so on down the line, he has arranged the papers in order of rank, or, in *rank order*. Having done this the teacher can immediately see where every pupil in the class stands in comparison with the other pupils.

When handling hundreds of papers, arranging them in this way becomes unwieldy. We can get the same information without having to handle the papers a great deal by filling out a frequency distribution sheet such as that given on the next page. Let us assume that a headmaster has given an informal arithmetic test to all classes of the middle school. He now wishes to find out where each class stands in reference to the others and also wants a standard by means of which he can quickly compare the standing of each pupil with the standing of others in his class. It makes little difference in what order he has his papers arranged, although for filling in the form it is best to separate the papers of each class as that will reduce the probability of entering the score of any paper in the wrong column. Let us assume that there were 60 possible points to be scored. We will group all those scoring from 60-56 together, those scoring from 55 to 51 together, and so on.

Let the examiner start with the papers of Class V. The first paper has a score of 43. A mark is then put in the column for Class V in the place allotted for those who get scores between 45 and 41. A second paper has a score of 27. A mark is put into the space allotted for scores between 30 and 26. When the scores of all papers in Class V have been filled in, those for Class VI are taken in hand and filled into the proper column. The examiner will save

himself some trouble of counting by arranging the marks in each space in groups of 5. This is conveniently done by putting down the first four marks thus////. When the fifth score of this classification comes along he does not add a vertical line but a horizontal line through the middle of the group of four, thus+///. This sets off the group of five conveniently

FREQUENCY DISTRIBUTION SHEET

Scores	Class V	Class VI	Class VII	Class VIII	Class IX	Class X
60-56						
55-51						
50-46						
45-41						
40-36						
35-31						
30-26						
25-21						
20-16						
15-11						
10- 5						
5- 1						
0						

After the scores of all papers have been entered, the marks may conveniently be changed into figures making a table like the following:—

FREQUENCY DISTRIBUTION SHEET¹

Scores	Class V	Class VI	Class VII	Class VIII	Class IX	Class X
60-56	0	0	0	4	0	3 ¹
55-51	0	0	5	11	27	54
50-46	2	4	20	17	19	21
45-41	6	11	19	35	35	14
40-36	12	22	30	53	17	1
35-31	30	40	49	45	16	1
30-26	51	55	29	21	15	1
25-21	62	38	23	12	3	0
20-16	49	19	8	3	4	0
15-11	15	7	5	0	0	0
10-5	13	0	0	0	0	0
5-1	0	1	0	0	0	0
0	0	0	0	0	0	0

How may the scores of Classes V and VI be conveniently compared? By averaging the score for each class separately. There are several ways of working out average values but the only two with which the amateur statistician need concern himself in the beginning are the

¹ The number in each column in heavy type denotes that the *median* is found in this frequency

median and the *mean*. All teachers commonly use these whether or not they know these terms.

The *median* can be assumed for practical reasons to be the *middle* score. This is the easiest of all the average values to compute and is for teachers perhaps the better of the two. If we have a stock of 100 papers arranged in proper order according to scores (in rank order) we need only count off 50 papers and the average of the score of the fiftieth and the fifty-first papers gives us the *median*. This is the most foolproof way of calculating an average score and is to be recommended for ordinary classroom use, or wherever the number of papers is not more than about 100.

Assuming, however, that we have many more papers, and have a distribution chart filled out like the model given on the previous page, we need only count the marks instead of the papers. If there are 480 marks, we will count half of these, 240, and the 240th mark will show us in which frequency the median score comes. If however, we have 'bunched' together all scores coming between say 55 and 51 we will want to know more precisely just where, between 55 and 51, the median comes. This can easily be calculated.

Supposing there are 20 marks or frequencies in the pigeon-hole for the scores between 55 and 51 and in counting our marks we have counted 230 before we got to the 55-51 group. We must then count ten of the twenty marks in this group to get our full 240. Since we used up just half of the marks in the pigeon-hole, our median score lies just half way between 55 and the top limit of the next group, 50. The median is then, 52.5. The ratio between the number of marks in the pigeon-hole and the number of marks that needed to be counted in order to find the middle score, gives us the clue as to where between

the top and bottom limits of the pigeon-hole the median score comes.

If the teacher who has his papers arranged in rank order wants to find out which 25 per cent of his students did best in the examination, he need only count off twenty-five per cent of the papers beginning with the papers with the highest scores. The names on the papers he has counted off give him his information. If he wants to find out the twenty-five per cent who did worst, he need only count off a quarter of the papers beginning from the end with the lowest scores. Similarly, if our scores are arranged on a distribution sheet we need only count off 25 per cent of the marks from the top of the distribution sheet and we are informed of the scores above which and below which the 25 per cent best and worst papers are to be found.

Exactly the same rule holds for working out these quarterway scores, or *quartiles*, as they are called in statistical parlance. They are to be worked out in the same way as the median except that a quarter of the marks on the distribution sheet are counted off instead of 50 per cent. In working out the median it makes no difference whether we count from the top or the bottom of the papers or distribution sheet, but in working out the quartiles, we must be certain to start at the high end for the best quarter of the papers, and from the low end for the poorest papers. Of course one may count off three-quarters from the high end to find the lower quartile. The term *quartile* is used extensively in statistics and should be remembered by the reader. It simply means the mark which shows the score above which or below which the twenty-five per cent best or worst performers are to be found. The highest quartile mark is designated as Q_3 (quartile 3) and the lowest quartile mark is Q_1

(quartile 1). There is also a Q_2 which is nothing more than the median or half-way mark. Sometimes one sees a Q written without a figure designating it. It means half the distance between Q_3 and Q_1 .

Percentiles are worked similarly to quartiles. We may want to pick out the best ten per cent rather than the best twenty-five per cent. We need only count ten per cent of the papers or marks beginning with the highest scores, and we have it. All those students who belong to the ten per cent which got the best marks are said to come into the ninetieth percentile because at least ninety per cent of the examinees got a lower score. The tenth percentiles are the ten per cent who had at least ninety per cent of the examinees doing better work than they.

The mean is worked out quite differently from the median. It is the average score. One way to find the average score on 20 papers is to add up the score of all 20 papers and divide by 20. Every teacher knows this method.

When hundreds of papers are involved, this method is rather tedious. If many scores or frequencies are in the same pigeon-hole of the distribution sheet, multiplication may be conveniently used. For instance, if there are 56 papers with a score of fifty and 63 papers with a score of 55, etc., multiplication will help to avoid part of the tedium of addition. One gets all the information needed to work out the average score from the distribution sheet. If there are 20 papers that scored between 55 and 51, the half-way mark of 52.5 is considered quite usable. Even with the help of averaging, or bunching the close scores in this way, and by use of multiplication, the process is a tedious one when dealing with large numbers of examinees.

There is a short method of calculating the mean which

is both easy and convenient though a little tricky to learn. Let us assume that 120 students make scores between 2 and 10 on a test according to the following frequency distribution sheet.

SCORE	FREQUENCIES
10	3
9	7
8	8
7	10
6	25
5	31
4	25
3	8
2	3

Glancing at this table we can immediately see that the average score must be about 5. Let us assume 5 to be the average score. If we guess wrong it does not matter for we are going to correct whatever error we make. The nearer right we guess the easier our computations will be but there is no harm done if we guess wrong. We then write a zero in the same line as the score 5 with its frequency of 31 (see table on next page). The meaning of this zero is 'no steps away from the assumed average or mean score'. The score of 6 is one step above our assumed mean so we write a one above it. The score of 4 is one step below the assumed mean so we write 'minus one' below the zero which indicates that it is less than average.

The other steps may be filled in similarly.

The amounts in the fourth column are obtained by multiplying the figures in the second and third columns.

Score	Frequencies (or number of pupils making the score)	Steps away from Assumed Score	Number of steps multiplied by frequencies
10	3	5	15
9	7	4	28
8	8	3	24
7	10	2	20
6	25	1	25
5	31	0	
4	25	-1	-25
3	8	-2	-16
2	3	-3	-9
Totals	120		62

Now add the plus quantities and minus quantities in the fourth column. This gives 'plus 62'. Divide this by the total number of frequencies (given in Column II) 62 divided by 120 gives 0.51—the amount of correction we have to make. Since in this case the amount of correction is a plus quantity we add this to the assumed mean which is 5. Thus our true mean is 5.51.

Supposing we guess 6 to be the mean instead of 5. Let the reader work out the result. The minus products will then be much higher than the plus products, which means that when we add them we have a minus quantity which is to be divided by the frequency (120). This will make the quotient also a minus quantity which again means in effect that we subtract instead of adding. Thus the assumed mean will be corrected by making it less than six and will work out to be the same as we secured by assuming the mean to be five.

The task is a little more complicated when we have grouped our scores. Instead of the range of scores lying

between two and ten, we will assume them to lie between 5 and 49. This time the assumed mean is half-way between 25 and 21 (see table below). In the score interval it is written 20-24 but what is actually meant is 20 to just short of 25, or some such figure as 20-24.9999999. Half-way between these two figures is approximately 22.5, which is our assumed mean. This is rather confusing to the novice. The easiest way to remember it is that if you start at 20 and allow an upward leeway of 5 points the average score is 22.5. When you come to make your correction you must again remember that the scores of the examinees did not go up one at a time in the frequency distribution but five at a time, such as 5-9, 10-14, etc. Therefore you must multiply whatever correction is to be made by five. A study of the last step of the process given above should make this clear. 22.5 (assumed mean) plus $(.51 \times 5)$ (the amount of the correction multiplied by the score interval which in this case is 5) equals 'true mean' 22.5 plus $(.51 \times 5)$ equals 22.5 plus 2.55 equals 25.5.

SCORE INTERVALS	FREQUENCIES	STEPS	FREQUENCY STEP
45-49	3	5	15
40-44	7	4	28
35-39	8	3	24
30-34	10	2	20
25-29	25	1	25
20-24	31	0	
15-19	25	-1	-25
10-14	8	-2	-16
5-9	3	-3	-9
TOTALS	120		62.0

Finding the mean is undoubtedly less fool-proof than finding the median. These two are not always approximately the same and therefore the statistician frequently demands to know both. But for general school purposes, the median alone or the mean alone are quite satisfactory and since the median is less confusing and less work to compute, the amateur is advised to use largely the median, unless he masters the above short method of computing the mean. Most 'norms' employ the median.

NORMAL CURVE

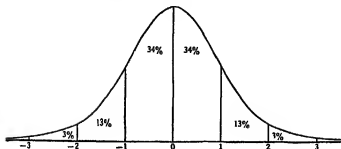
A statistician is by no means satisfied with knowing only average or median scores. He wants to know whether the rest of the scores cluster around the mean or median or whether they distribute themselves all over the possible range.

First he will want to know the range of the testing. This is the distance between the top score and the bottom score. If in a test in which it is possible to score 100 marks the range is 75-100, that is, all scores fall between 75 and 100, we immediately know that the test was too easy. If the range is 0 to 37 we know the test was too hard for the group. If the range was from 10-95 we know that the test has probably been hard and easy enough to give all a fair chance, but not necessarily so. If 75 per cent of the examinees got between 80 and 95 and the remaining 25 per cent got between 10 and 80 the test was obviously lopsidedly easy. Likewise it was lopsidedly hard if 75 per cent scored from 10 to 20 and a few got numbers all along the range up to 95. If on the other hand, in the range of 10 to 95, about one-half got from 40 to 60 and the frequencies gradually taper off in number towards the highest mark and the lowest mark, we have a well balanced test that divides the examinees off according to

natural abilities. If we show these results in graph form we will approximately have what is called a normal probability curve

A graph is called a normal probability curve, or is said to approximate the normal curve, when it shows the same results as many other measurements in human abilities or traits (if the people who are examined are a random sampling and not a selected group). One does not find that all people are either very good or very poor in any trait or ability. Most people are nearly average. Sixty-eight per cent of the people usually get a score in a reliable measurement that gives a fair chance to all, in that one-third part of the range which is cut in half by the average score. Only twenty-six per cent usually get scores in the second sixth and the fifth sixth of the range. Whether one measures intelligence, or height, or weight, or the size of feet, or arithmetical ability, or running ability, one finds that people are divided into very similar classifications. Even in throwing dice or tossing coins tabulations of the results in a graph will give us a curve very like the illustration on opposite page. This is a most interesting phenomenon which many try to explain but no one has explained satisfactorily. The significant part is that it is so often true. Sixty-eight per cent of us are within one-sixth distance of being average in any trait or ability. A table similar to that on page 161 on the distribution of intelligence could be repeated for almost everything we do

So much faith is placed in the probability that most measurements will reveal a similar result, that any test which does not show results that can be plotted in a graph something like this curve, is suspected of not being natural, and the group tested is suspected of not being a random sample of the population. Most curves are expected to resemble the normal one approximately. Some



The straight line on the bottom is divided into six equal parts, signifying that the range of the measurement has been divided into six equal parts. The middle division line indicates the average score or measurement. Normally 34 per cent of the people tested or measured will get a score in the division immediately below the average score and 34 per cent will get a score in the division immediately above the average score (in No. -1 and No. 1). 13 per cent of the people will get a score in No. -2 and in No. 2. Only about 3 per cent will get a score in No. 3 and in No. -3. This means that.

68 per cent score within one-sixth distance of the average score (in the two middle divisions).

26 per cent score in the second and fifth divisions (Nos. 2 and -2) of the range.

6 per cent get extreme scores (Nos. 3 and -3).

Only a perfect type of measurement and an absolutely normal sampling of people can be expected to show exactly this normal probability curve, but a rough approximation to this curve will be found in most measurements if the measurement is sufficiently well graduated.

graphs look like the outline of jagged mountain tops and are not bell shaped. But nevertheless the peaks build up on a huge mass which indicates an approximation to the normal curve. Our ways of measuring are imperfect and these imperfections are in part revealed in the jaggedness of our graphs. This does not say that if the results of a certain measurement do not approximate to the normal

curve there is necessarily something wrong with the measurements. One cannot be certain that the results of the measurements of all abilities and traits will approximate to those indicated by the curve until one proves it experimentally; statisticians only say that very many measurements do and it is worth inquiring into the reason when they do not.

A normal curve is not to be expected when testing only a few cases. At least one thousand measurements are needed and even this is a small sampling. A sampling of 100,000 is desired in some cases.

STANDARD DEVIATION

It is not surprising that the statistician, finding that in most of our measurements the results should approximate to a normal curve, has made a mathematical device which tells him the range of the score within which, approximately, the sixty-eight per cent of the scores that are closest to average come. This device is called the *standard deviation*. When the standard deviation of a certain measurement is 10, it means that within ten scores on either side of the mean (a range of 20) sixty-eight per cent of the frequencies are to be found, or, what is the same thing, the score-limits within which sixty-eight per cent of the examinees got their score, is 20 points. With this information in hand the statistician can make a rough guess as to how the other thirty-two per cent of the examinees fared on this measurement provided the examinees were distributed over the range as in a normal curve (which is to be expected but not to be depended upon). Thus the standard deviation is a measurement of the *spread* of the scores over the range of the scores. If in addition to the standard deviation (which is represented by the Greek letter sigma or the initials S.D.) the range

and the mean are indicated, the statistician has a fairly accurate picture in his mind of what the measurement found.

The standard deviation has a relative value. It means a great deal to the statistician who knows approximately what standard deviation is to be expected on each particular measurement for examinees of certain ages. Older examinees are expected to have a larger range in ability and therefore a higher S.D. in the measurements made upon them. The stiffness of a measurement also helps to determine the range. There are a number of considerations which, unless one has standard deviation data on other similar measurements conducted on similar examinees, make an unrelated S.D. figure useless. Unless one knows approximately what S.D. is to be expected under similar measurements and circumstances, an isolated S.D. figure is completely useless. So although the S.D. means much to the statistician it is not likely to mean much to the amateur. But the amateur should know that it is a mathematical device which shows whether most of the examinees made nearly the average score or whether the scores were scattered over a wide range. References to the S.D. are so numerous in literature on modern tests that the amateur needs to know what it is about.

It is not very difficult to compute the S.D.; but to go into mathematics in sufficient detail to help to a real understanding of complicated processes like S.D., correlations, and probable error is beyond the purpose of this book. What can be given in one chapter is not sufficient to be of real help. He who wishes to go in for statistics will need a more adequate guide such as the two recommended in the beginning of this chapter.

Overleaf are given the S.D.s for several columns of figures. Note the change in character of the data in each

column as the S.D. goes up. No. 1 tells us that seventy per cent of the scores are found within 5 points of either side of the mean or quite close to the mean. No. 4 tells us that most of the scores were made relatively far away from the mean. No. 1 approximates a normal curve while No. 4 is about as far from normal as it can be.

SCORE	FREQUEN- CIES ON TEST 1	FREQUEN- CIES ON TEST 2	FREQUEN- CIES ON TEST 3	FREQUEN- CIES ON TEST 4
45-49	0	40	18	30
40-44	0	22	12	15
35-39	15	30	12	10
30-34	30	14	12	5
25-29	50	15	12	0
20-24	20	5	12	5
15-19	15	0	12	10
10-14	0	0	12	15
5-9	0	0	18	30
S.D.	5.5	7.85	10.5	16.4

CORRELATIONS

One cannot delve far into educational psychology without using a very handy device called the 'coefficient of correlation'. This technical device is used by several sciences including astronomy, physics, biology, psychology, economics, engineering, and education. What it tells us is the relationship between two or even more factors. Almost anybody who reflects at all on the causes of the phenomena we see about us makes correlations in at least a haphazard way.

For instance, when we expect the older of two unknown schoolboys to be the taller of the two and we also expect

him to be in a higher class in school, we show that we have an understanding of the co-relation (or correlation, as it is usually written and pronounced) between age and height, and between age and scholarship. That does not mean that we invariably expect the older to be taller and more advanced in school, but as a rule that is what we do expect. When we say that the bright children have better results in school than the dull, we mean that there is a definite correlation between brightness and scholarship.

When two factors correlate they may be in the relationship of cause and effect. But again they may not. There is a correlation between height and weight but height is neither the cause nor effect of weight, nor weight of height. Rather both are the result of a variety of causes such as diet, state of health, heredity, age, etc. Correlation is no proof whatever of cause and effect but simply a statement that when you have the one, you are likely to have the other also.

An educationist is in constant need of knowing the answer to certain questions involving correlation. He will never know what kind of education he is imparting until he knows the answer. For instance, he needs to ask: 'Are those who read the most rapidly the ones who understand best what they are reading? Are those who are the best in mathematics the ones who can best work out the kind of problems they will meet in the practical world? Are those who are weak in reading the ones who are weak in the ability to make themselves understood?' Questions of this nature can, since a scientific procedure in the examination of educational matters was developed, be answered with a fair degree of accuracy. In the past such questions were answered by asking the opinion of individuals, which is another name for 'guesswork'. Nowadays there is a

scientific approach which, when applied discriminately, can give an answer which is not far from the mark

Let us take, for instance, 100 children and examine them with a carefully worked out geography test and then with an arithmetic test. Let us then put down the name of the examinee who got the highest number in geography first, the second highest next, and so on down the line with the entire 100, ranking them according to results in arithmetic in the same way. Now if these two lists are exactly the same and every pupil has exactly the same standing, relative to his fellow examinees, in both tests, there is a perfect correlation between the abilities measured by the test. Of course there will not be such a result. 'In education a perfect correlation is practically impossible. Most of the pupils who are good in arithmetic will be good in geography, but there will be many cases of being good in one and weak in the other. We can measure the amount of correlation there is. We can even measure negative correlation, that is, tell whether strength in one subject is usually accompanied by weakness in another. There are degrees of correlation between two subjects that are as easily, though not quite as reliably compared as the degrees of attraction and repulsion of the different parts of a magnet.

In fact an illustration of the mathematical designation of the amount of correlation looks like a bar magnet with its positive and negative ends, the positive force weakening as we travel towards the middle where there is complete neutralization. Negative quality is found towards the other end. At the positive pole we have the symbol 'one' and at the negative pole the symbol 'minus one'. Here is the way one might picture it:

1.00	0.75	0.50	0.25	00	-0.25	-0.50	-0.75	-1.00
------	------	------	------	----	-------	-------	-------	-------

1.00 means that two traits or skills or factors correlate perfectly. 0.75 means they are strongly alike to three-quarters extent. 0.50 means there is a decided correspondence between them but only half as strong as perfect correspondence. 0.25 means there is correlation of quarter strength. Zero correlation means pure chance, nothing more than a chance relation between the two. A correlation of -0.25 means, the less you have of the one the more you will have of the other. -0.50 means the same to a stronger degree. -1.00 means exact opposites like heat and cold, honesty and dishonesty, etc., if indeed these be quite the opposite of each other.¹

A correlation of 0.50 must not be interpreted as meaning that the chance is 50-50 that he who is good in arithmetic is also good in geography. A 50-50 chance is zero correlation. A correlation of 0.50 is half way between a perfect correlation and 50-50 chance, which means a quite decided correlation, almost enough to say, 'as a rule'. The coefficient of correlation must not be confused with percentage. If we say two Intelligence tests have a correlation of 0.50 we do not mean that they are 50 per cent accurate. We mean that these two tests measure the same thing to the extent that is about half what it would be were they identical. (Incidentally, in education there are no tests

¹ The above explanation is, as any very concise explanation must be, an over-simplification and must be regarded as only approximately accurate. The amount of correlation is expressed mathematically which gives an expression of exactness which does not quite correspond to fact. We cannot say that the difference between .80 and .90 of correlation represents exactly the same amount of correlation as the difference between .10 and .20, as we can say that a difference of ten degrees of temperature is exactly the same difference in amount of heat whether near freezing-point or near boiling-point. The amount of correlation is not nearly as definitely fixed a quantity as the mathematical expression of it seems to suggest.

that agree even with themselves 100 per cent of the time. Any test which can be given twice and have the results correlate between .80 and .90 is considered a fairly reliable test. Mental abilities cannot be measured as accurately and reliably as the expansion of a rod of iron under changes of temperature.)

The mathematics of working out correlations is rather intricate and will probably be used by hardly anyone who is not doing serious research work. The theory and mathematics of correlations cannot be treated in this book for lack of space. But understanding what is meant by a correlation is essential for anyone who reads progressive educational literature, because the facts on which the changes of modern education are based owe much to what was learned in using the technique of making correlations, and these facts are continuously being expressed in the terms explained above.

The amount of correlation depends to some extent on whom we are examining. If we chose eighteen boys with birthdays one year apart, it is possible that we should find a perfect correlation between their heights and weights. At any rate we would find a very high correlation. If on the other hand, we took eighteen boys all six years of age, we would find that the correlation between their weights and heights was much less. This is easy to understand because among the eighteen boys of different ages the differences are so great in both height and weight that the rule 'the taller the heavier' is obvious by merely looking at the boys. Amongst the six-year-olds the rule is not so obvious and although the rule still holds, it does not work quite so smoothly. Similarly, if we were to correlate the ability to read and to do mathematics of boys of each class from kindergarten to college, we would find two mental abilities correlating

much more highly than if we examined the pupils of Class V where those examined are much more nearly of the same ability. We must expect much higher correlations with examinees of unlike ability than with those of nearly the same ability. A correlation of 0.25 in two abilities among pupils of Class I may be more significant than a correlation of 0.50 among pupils of Classes I to V. A correlation is precise only when it is clear what group was examined.

Then, too, the correlation depends partly on the test. A test in which the pupils score anywhere from 10 to 100 questions right is much more accurate than one in which all the pupils scored between 90 and 100 or between 10 and 20 although there were 100 points to be made. The first test really separates the weak pupils from the strong while the last two are either so easy or so hard that the differences in the abilities of pupils were not found out at all. It takes a good test with many items to be useful in finding the correlation of two abilities. Generally speaking, it takes a Standardized test or at least a skilful objective test to be of use. The ordinary essay type of examination in which different examiners give widely divergent marks is useless. The name or description of the test used is an essential part of the information regarding the correlation.

The correlation device is used extensively in research work. One cannot read much on the scientific aspects of education or psychology without finding references to it. In this book the author has been practically forced to refer to it a number of times in reference to the computation of the reliability and validity of tests, in discussion of how prognostic tests are composed, and in reference to the reliability of the results of Old-type tests. Educational psychology has found the method of correlation to be one

of its most usable tools in working out what skills or traits or abilities may be considered related.

One usually finds the coefficient of correlation written in the following way: -0.80 ± 0.02 (0.80 plus or minus 0.02). The latter figure denotes the *probable error*, a statistical check to see how reliable the coefficient of correlation is. It means that there is a fifty per cent probability that whatever error may be in the figure given for correlation, due to chance elements, will not exceed 0.02; in other words, there is an even chance that the true figure, freed of all chance elements, will be somewhere between 0.82 and 0.78. The amateur need not concern himself about the probable error. Its value is to the statistician and research worker and may be ignored by others.

If only a few examinees are examined, the probable error will be large; if many are examined it will be smaller, for the more examined the less the probability that these particular scores were freakish.

One may also compute the probable error of the mean, that is, give the figures which show the probable limits of error in statistical averaging.

The *standard error* is similar to the probable error and indicates a two-to-one chance that the true mean will be found within the limits indicated instead of the even chance designated by the latter.

XI

CONCLUSION

THIS book will be concluded with a brief appraisal of the various types of tests described and discussed and a note on their probable usefulness in India in the next few years, provided proper effort is made to use them. Proper effort does not mean effort to duplicate the rather extravagant development taking place in America, but modest effort to take advantage of some of the undoubted advantages of scientific measurement in education and psychology.

In America the development of Standardized tests has not been a mushroom growth, as is often alleged, although extravagant developments and exaggerated claims have been made for them. The entire school system in America has become interested in Standardized tests and it is not necessarily a mushroom growth because several million pupils use several million tests a year. Let it not be forgotten that it would also take as many millions of sheets of blank paper a year to examine such a large number of pupils in several subjects by the traditional methods. The cost of printing in India need only be about 25 per cent more than that of blank paper. Objective tests have not proved economically extravagant and have helped to put education on a more scientific footing.

In the British Isles the use of such tests has not been on the mass scale found in America but the movement met with the favour of progressive leaders from the very

beginning as is shown by the development in Northumberland and Fife.

Which types of tests show promise of general usefulness in India? Which types are undoubtedly practical and which types are still in the experimental stage? Which types are usable by the expert alone and which are usable by any average progressive teacher? Attempt will be made to answer these questions.

Of those types still distinctly in the experimental stage, prognosis, aptitude, and many of the psychological tests (such as evaluating the emotions, temperament, conduct, etc.) must be mentioned. Intelligence tests have in many respects passed from the experimental stage into the stage of positive usefulness, but there are many points respecting the use of such tests which are still in the experimental stage, especially as regards their use among primitive or semi-primitive peoples. Revision of opinion as to just how much one can depend on the findings of Intelligence tests is still going on among the experts.

When we ask which tests need to be administered and interpreted by experts, we must mention the same groups. The various Mechanical Aptitude and Industrial Aptitude tests are not to be handled or interpreted by any but the professional; so also all but a few of the distinctly psychological tests. Prognosis tests are too new as yet for one to know whether even the professional can do anything with them. The same must be said for some of the Aptitude and Psychological tests. The results with all these tests is encouraging and worth experimenting with in India, but they are not practical for the amateur at present.

Intelligence tests of the performance type require a carefully trained examiner to administer but there are no few 'pencil and paper' tests which are objective and

simple enough for a discriminate layman, who will carefully observe the directions, to give. As to actually making use of the results of such tests there is considerable difference of opinion. It is a moot question whether it is at all wise for the teacher to know the intelligence quotient of the pupil. A teacher with such knowledge may take the results too seriously, quickly interpreting every failure on the part of certain pupils to be due to low intelligence and giving unfair preference in his attentions to those who are supposed to be bright. Some teachers cannot help but let this information prejudice them in the treatment of pupils so that the backward pupil senses the teacher's low estimate of his abilities and develops an inferiority complex, and the pupil with a high IQ on the other hand is encouraged to assume an attitude of superiority. In schools where all pupils are given Intelligence tests as a matter of routine the information thus gained is usually accessible only to the principal and psychological adviser and not to individual teachers.

Only the exceptional teacher is able to use such information wisely. Therefore only the exceptionally well-informed teacher should be entrusted with the information except as it occasionally sheds light on a particular case. The information gathered from Intelligence tests as it applies to individuals is to be treated as strictly confidential.

There is also the question as to whether the examinee should be told his IQ. Those of low IQ usually have little understanding of what it means after they have been told. Those of average intelligence may be disappointed in finding out that they are not superior, but it is also essential for us to know our limitations and to realize that if we wish to succeed we must be prepared to work somewhat harder than our neighbour with a higher IQ. No rule can

be given as to how much an examinee should be told of his own condition. As a physician uses discretion in revealing to his patient his condition, so the psychologist must use discretion. Much depends upon the personal relationship between examinee and his examiner or adviser.

In India where only a small percentage of the population can receive education above the primary standard and where the economic pressure has been so great that relatives goad boys on because of improved pay prospects, even though they are incapable of satisfactory work and are over age, Intelligence tests are needed to select the small percentage that shall have the privileged opportunities, and also to separate the mere rote-memory learners from those of natural ability. Intelligence testing will be the province of the principal, special adviser of the students, or an experimentalist, rather than of the individual teacher. At present the greatest need is for good Intelligence tests for the educational experimentalist, as many of the problems of administration, curriculum, and selection of pupils need to be studied with reliable information in hand as to the natural ability and possible development of students. In this respect Intelligence tests can supply much of it.

The type of Standardized tests which are of greatest use to teachers are the Scholastic tests. These any intelligent teacher can use. The use of such tests is no longer in the experimental stage. Their usefulness has been proved and they can readily be applied. The individual test is an experiment until it can be proved that the test does what a good arithmetic or reading or science test is supposed to do: but there is no longer doubt about really useful tests being possible.

It is in this field where there is so little doubt about

valuable help being available that so little experiment and adaptation has been done in India. There has been more experiment with Intelligence tests which cannot be put to use so readily by the ordinary teacher than with Reading or Arithmetic tests which can be put to immediate use whenever and wherever wanted by any capable teacher. The Intelligence test can only give information about a hopeless situation. (The situation is termed hopeless because whatever intelligence a child has, it has. The teacher cannot add greatly to it or subtract greatly from it. Nature's gift of more or less potential intelligence has been given to us for better or for worse and is not exchangeable.) Scholastic tests, however, reveal weaknesses in the pupil that can very often be remedied, and strengths that can be made into foundations for larger constructions. The Intelligence test reveals little to the teacher of his own successes and failures but a Reading test may give a sharply focussed profile of the teacher as teacher.

Widespread use of good Subject or Achievement tests could have a most profound influence on both teaching and curriculum, not only on the method of teaching and the contents of the curriculum but on the spirit and aims of teaching. It would also aid in the much needed reform of the examination system in India and provide definite standards for achievement in place of the present muddled standards. But hitherto very little has been done in India in this most practical aspect of the entire movement towards scientific measurement in education.

As regards the possible use of informal New-Type tests by the teacher, nothing further need be said. These are evidently usable by anyone who learns to construct them. Achievement tests supply no end of suggestions as to proper models. Much of the benefits of the Standardized

Achievement test can be got from informal tests¹ constructed on a modest scale along the same lines.¹

Diagnostic tests, as mentioned before, can be used by any teacher who will examine the work of his pupils with these questions in mind: 'Why does this particular pupil have difficulty in doing this piece of work? Just what does he understand of the process and what does he not understand? What is the defect in his skill or where does his reasoning go wrong?' Any test can become a useful diagnostic test for the teacher who realizes that every individual has his peculiar difficulties which the teacher must locate if he is to help the pupil. Some tests show up the difficulties of pupils more than others and a good test especially constructed for this purpose is considerably more helpful than an ordinary test, but more depends on the inquiring mind of the teacher than upon the form of the test.

In conclusion, the modern testing movement is more than just a flood of interesting novelties. It is a way of measurement that greatly refines the crude methods of measurement employed in the educational and psychological field heretofore. The tape measure and the square add much to the precision of a carpenter's work. Similarly modern testing methods, when used discriminatingly, add precision and standards to the educationist's work. There is nothing that has added as much to educational advance in America and England in the last generation as the better technique of measurement. It has been said that

¹ A noteworthy attempt has been made by Mr D'Sylva and his team of officers on special duty for the Department of Education of the Central Provinces to introduce improved informal testing methods on a modest scale into the ordinary primary school. See *Measurement of Attainments in Primary Schools*, published by the Government Printing Office, Nagpur, 1937.

the only original advance in education of the last hundred years is the refinement of measurement. India will not progress toward a progressive and up-to-date educational system unless she improves her way of educational measurement. The modern testing movement has much to offer and probably no other one thing will do as much to break Indian education away from the fetters of its admittedly legalistic, stereotyped, and examination-dominated education (which is neither oriental nor western) and launch it on a course of open-minded but safely-piloted experiment which will bring to her a more progressive and at the same time more naturally Indian education. With a better way of measuring educational efforts India would be freer to experiment and could much sooner reject that of which the value is doubtful. As it is at present many questionable practices hang on interminably because no one checks their results, which in many (but by no means all) cases are easily ascertainable with the means of measurement now at our disposal.

APPENDIX I

YEARLY TEST FOR CLASS III
OF THE PRIMARY SCHOOL

THE following test was composed by Mr B. Sadhu, Supervisor of Primary Schools, Bismampur, C.P. This test is for Class III and aims at covering most of the work required in the syllabus of this class. A few of the items are to be administered individually or orally. In the case of handwork and physical culture, the work of the pupil must be observed. But for the most part all that the pupil has to do is done in the printed examination paper.

This examination takes about three two-hour-periods to administer. Parts of it may be omitted without prejudice to the test as a whole. It is not a fully standardized test as norms are not supplied. It was printed for use in fifty schools and was supplied as a 16-page folder to all pupils at one anna a copy, which shows how cheaply extensive testing can be when a number of schools club together.

The test is included here as an illustration of a test using both old-type and new-type items which tests most of a year's work in a comparatively short time.

VOCABULARY TEST

1 point for each right answer [20 minutes]

1. *Write the meanings of these words:—*

Pleased

Western

Difficult

Underline the synonym of the first word in each line:—

Suddenly	(Yesterday, quickly, without warning, repeatedly, wonderfully)
Praise	(Compliment, hymn, scold, sing)
Leader	(Master, soldier, police, one who goes before)
Accept	(Confess, deceive, receive, except)
Attempt	(Finish, succeed, fail, try)
Discourage	(Dishearten, disgrace, defeat, desert)
Debate	(Quarrel, argue, oration, declare)
Anger	(Laughter, weeping, violent, rage)

2. Write the word with the opposite meaning of the words given below:—

Swiftly—	Soft—	Expensive
Finished.	Just	Cruel

Underline the word with meaning opposite to that of the first word in each line.—

Profit	(Shop, loss, sale, rupees)
Master	(Chowkidar, kotwal, servant, coolie)
Strong	(Weak, healthy, defeated, gentle)
Front	(side, back, south, answer)
Outside	(Palace, house, inside, well)
Medicine	(Liquid, bitter, sweet, poison)
Humility	(stubborn, unjust, pride, yield)
Fresh	(Green, stale, worthless, bold)

3. Write a sentence with each of the following words:—

Power of hearing	_____
Cavalry	_____
Attack	_____
Earning	_____

4. Correct the spelling of the words:—

Bhagwan..	Niyayt.	Dud
Apradhu	Kijiyega	Karay

5. The examiner will select 25 isolated words used in the readers and give them as a spelling lesson. 12 points if all are right. 1 off for each mistake down to zero.

READING (STORY)

Read Silently:

[10 minutes]

A merchant wished to go to Bombay and so went to the railway station in plenty of time. When he arrived he realized he had forgotten his trunk. So he quickly jumped into the tonga again and drove home. The tongawalla made his horse go very fast. When they arrived back at the station with the trunk the train was already at the station. The merchant paid the tongawalla and called a coolie to carry his baggage. No coolie came. The trunk was very heavy for it had Rs. 3000 in silver in it. At last the merchant seized the trunk himself and ran to the ticket office to get his ticket. The ticket babu said 'I am sorry. It is late. You will miss the train. Tell the guard that you have no ticket. He will sell you one at the next station'. The merchant ran towards the train but the train started. The guard was very kind and stopped the train. The merchant lifted his trunk with 37½ seers of silver in it on the train and climbed aboard. He was all out of breath and very hot and tired when the train started. He fanned himself a long time before he felt comfortable again.

Some of the statements below are true, and some are false. Underline 'true' if the statement is true, and 'false' if it is false.

The merchant wished to go to Bombay	<i>True</i>	<i>False</i>
After the train had left he remembered		
his trunk	<i>True</i>	<i>False</i>
He went to the station in a horse tonga	<i>True</i>	<i>False</i>
The tongawalla drove very fast	<i>True</i>	<i>False</i>
The train left before the merchant		
arrived with his trunk	<i>True</i>	<i>False</i>
The merchant did not pay the		
tongawalla	<i>True</i>	<i>False</i>

A coolie came as soon as the merchant		
called	<i>True</i>	<i>False</i>
The trunk was not heavy	<i>True</i>	<i>False</i>
The ticket babu quickly sold him a		
ticket	<i>True</i>	<i>False</i>
There was Rs. 3000 in the trunk	<i>True</i>	<i>False</i>
The money was in notes	<i>True</i>	<i>False</i>
The ticket babu told him to tell the		
guard he had no ticket	<i>True</i>	<i>False</i>
Rs. 3000 in silver weighs $37\frac{1}{2}$ seers	<i>True</i>	<i>False</i>
The guard was helpful	<i>True</i>	<i>False</i>
The guard told him not to get on the		
train without a ticket	<i>True</i>	<i>False</i>
The guard stopped the train	<i>True</i>	<i>False</i>
The merchant caught the train with		
difficulty	<i>True</i>	<i>False</i>
He missed the train because he could		
not lift the trunk	<i>True</i>	<i>False</i>
He was very hot and breathless	<i>True</i>	<i>False</i>
He could not get a fan to cool himself		
off	<i>True</i>	<i>False</i>

2 points for each right answer

2 points penalty for each wrong answer

POETRY

What is the meaning of the 2 verses below?— [8 points.

1. Bahut n shit, n ati garmi hai,
Kahin ann-jal ki n maki hai.
Vayu yahan ki hai sukhkari,
Rog nahin hai koi bhar.
2. Nirjan ban men giri se girkar,
Bahta rahta hai kal-kal kar.
Kitna swachchh aur shital hai,
Chandi ki dhara si jal hain.

Read the poetry given below, then answer the questions:— [8 points:

Vyarth widesh nahin main jata,
Jata to kuchh lekar ata,

Pakar mujhe tumhen sukh hota,
Agar kahin main rupya hota.

Main nirarthak kahin kahan nahin jata?

Kise pakar tumko sukh milta? ...

Widesh se rupya kya lata?

In panktiyon ko kaun kahta hai?

COMPOSITION

Write the story of 'How our Class made a Garden'.

[40 points]

ORAL READING

1. The teacher will assign a paragraph. 5 marks may be given for pronunciation and 5 numbers for each ten words read per minute.

Follow directions. Read silently.

2. Do as you are instructed. (Time. 3 minutes)

[20 points]



Draw a circle round the elephant.
Colour the cow black.

- Write Gandhiji's name under his picture.
 Draw four lines under the Charkha.
 Write a G under the gramophone.
 Put an arrow at the part of watch used to wind a watch with.
 Colour the glass in the spectacles black.
 Write a 5 in the middle of the star.
 Write your name under the lantern.
 Put a cross under the heel of the shoe.

ARITHMETIC

- 1 (a). Write these numbers in order according to their ascending value. (Time: 5 minutes) [6 points]

775, 905, 100, 567, 899, 612, 325, 360.

- (b). Subtracting 5 each time start with the number 67 and continue to 27:—

67, —, —, —, —, —, —, —, 27.

2. Write these in value numerals:—

(Time. 5 minutes) [6 points]

Twenty-five thousand fifty

Three lacs five

Thirty-nine thousand one

Fifty-nine lacs three thousand

Three hundred sixty-seven

Eighty-nine

3. Addition. (Time: 20 minutes) [16 points]

	Rs.	ps.	as.	Md.	sr.	oh.
40567	9	12	9	47	37	6
4102	10	9	6	108	28	15
645	8	15	6	8	10	12
250						

Subtract.

	Md.	sr.	ch.	Yds	girah	Tola	masha
61058	88	12	13	80	12	18	5
9562	37	28	14	28	14	12	9

Multiply

	Rs	as.	ps	Rs.	as	ps.
6758	12	8	6	26	8	12
523			× 28			× 54

Divide

	Rs.	as.	ps.
82) 6456 (20) 646	12	6(76) 48240(

4. *Write the Answers of these:***[12 points***(Time. 6 minutes)*

120 inches	= ft.
90 inches	= ft.
64 annas	= Rs.
32 pice	= annas
48 furlongs	= miles
84 masha	= tolas
63 inches	= . . . yds. . . ft. . . in.
75 pice	= . . . Rs. . . . as. . . ps.
57 seers	= . . . mds. . . seers
15 paili	= . . . kathas . . paili
138 ratti	= . . . tola . . masha . . ratti.
99 chhatak	= . . . seers . . chhatak
360 minutes	= hours
80 quires	= reams
120 days	= months
36 weeks	= months
84 months	= years
2 sheets	= quires

ARITHMETIC REASONING

Write the answers of these 9 problems in 20 minutes.

(2 numbers each for these six problems)

1. How many pao in $1\frac{1}{2}$ seers? ()
2. At as $2/6$ per pao what does 1 seer of jalebies cost? ()
3. Mohan weighed out 2 maunds 10 seers of dal. How much must he add to make it $2\frac{1}{2}$ maunds? ()
4. It is now 4.30. What time will it be after 45 minutes? ()
5. 12 umbrellas cost Rs. 18/-. What is the cost of 7 umbrellas? ()
6. I went to the bazaar with Re. 1/-. I bought, at the rate of — per seer, 3 seers of potatoes, — worth of sugar and — of chillies. How much money should I have left? ()

(4 points for each right of the following 3 problems)

Solve these problems showing clearly the method.

1. 16 blankets cost Rs. 80/8/0. What will 12 cost?
2. A merchant bought 30 umbrellas at Rs. 61/14/- and sold them at Rs. 12/2/- each. What was his profit?
3. I went to the bazaar with Rs. 27/-. I bought a dhoti for Rs. 3/8/-, cloth for Rs. 6/10/- and a blanket for Rs. 13/10/-. How much money should I have left?

HYGIENE AND PHYSICAL CULTURE

1. *Fill in the blank spaces:* [6 points]
 The mouth should be..... while breathing.
 It is best to play where the air is ..
 By drinking unboiled milk we run the risk of catching
 .. — . . . disease. disease is
 spread through flies.

2. *Write the answers of:* [8 points]
 What time of day is best for sweeping the school?

 What harm is there in keeping windows closed tight
 at night?
 What harm is there in smoking for children? ..
 What harm is there in drinking intoxicants? (Name 2)
 (i) (ii)
 3. *Observe posture, discipline, cooperation in and
 observing the rules in games, cooperation in march-
 ing, and keenness.* [25 points]

GEOGRAPHY

1. Draw a map of your district and show these places
 on the map.

Your village, the road from your tahsil headquarters
 to your district headquarters, 3 prominent villages
 on the road, 2 places where there is a bazaar, the
 place of a mela, the biggest river in the district, a
 mountain range (if any), the section where winter
 crops are grown, railroads (if any). [20 points]

2. *Answer these questions:* [20 points]
 a. At what time of the year are 'Unhari' crops
 planted?
 b. What kind of soil is best for building mud homes?

 c. What is the monsoon?
 d. What tahsil of your district is the most heavily
 populated?
 e. What are the principal industries of your district?

 f. What kind of fresh vegetables come to the bazaar
 in July?
 g. By the altitude of which river is the general level
 of your district calculated?

- h. What body is responsible for the primary schools of your district?
- i. What is the principal forest product of your district?
- j. From which direction does the wind come mostly during the monsoon in your village?

CITIZENSHIP

[80 minutes]

1. What is 'safety first'?
- What are the principal handicrafts in our district?

What products are got from the forests along the biggest river of your district?

(Name 3 kinds of wood and 3 other products)

What differences do you find in a railway station and a motor bus stand? Give 3 differences. [11 points]

2. Underline the word true if the statement is true, and false if the statement is false:— [11 points]

When a master comes into the school we should remain sitting. *True False.*

When a guest from afar, comes to your house, his name should be sent to the Mukaddam. *True False.*

The river and the canal do the same work. *True False.*

The malguzar may send a man to jail. *True False.*

In order to get one's mail from the post office, one should give the postmaster a baksheesh. *True False.*

Only officials may send goods on the railway. *True False.*

You are entitled to open the mail bag if the postmaster is slow about it. *True False.*

The tahsildar may buy any fields he likes for himself. *True False.*

The patwari is next in rank to the tahsildar. *True False.*

The patwari is under the rank of the revenue inspectors. *True False.*

Railway tickets may be bought at the post office. *True False.*

3. Write the words of the song 'My River'

[6 points]

4. Make a model of a railway station in the sand showing signals, etc. using paper, cardboard, bamboo, etc. to complete the model. (Time 40 minutes). [12 points]

HANDWORK

1. Make a hollow mango out of earth This may be baked later.

2. Practical work:—

Colour the mango.

Make 1 yard of rope.

Make a bhindi and a plantain out of earth.

Make 2 flowers out of coloured paper.

[40 points]

APPENDIX II

THE MIDDLE SCHOOL ACHIEVEMENT TEST (IN HINDI)

THE purpose of this test is quite different from that of the test in Appendix I. That test attempts to review nearly the whole syllabus for a year's work. This test very largely ignores the syllabus of the past year and conducts a rapid survey of the work done during the last five years or so. It was designed as a High School or Normal School entrance test to survey the educational foundation of the pupil before admitting him into a higher institution. It is a rapid review of the fundamental skills and knowledges.

It has nine sub-divisions, grouped in two parts. Part I contains Hindi reading, arithmetic computation, arithmetic reasoning, and English. Part II contains vocabulary, grammar, history and civics, geography, science and hygiene. Each part takes only two hours to administer so that the nine subjects may be examined in four hours. The reliability is .83. Part I alone has a reliability of .81 so may be used alone without part II. This test does not pretend to be thorough in every subject (for that it would have to be lengthened) but it is reliable enough in each subject to give at least a fair indication of the pupil's ability and when the results in all subjects are used to determine the all-round ability of the pupil, may be considered to give a highly reliable profile, even when only Part I is used. The test is designed for reliability in all-round ability rather than ability in each subject separately.

For reasons of space, the entire test cannot be reprinted but the extracts sufficiently indicate what it is like. The omitted test in Hindi reading is much like the one here printed in English but is more difficult because Hindi is assumed to be the mother-tongue while English is a

foreign tongue. The tests in history, geography, and grammar are much like the included example on science.

The norms (both the median and the quartiles) based on 3000 tests are included. These norms show the achievement in all subjects separately as well as in aggregate. The test was composed by the author in collaboration with the staff of Bhatapara Middle School.

ARITHMETIC COMPUTATION

Time 30 minutes

Note. Reduce fractions in answers to the lowest denominators.

- | (1) | $\begin{array}{r} 29788 \\ 42391 \\ 2430 \\ 36875 \\ 51134 \\ + 777 \\ \hline \end{array}$ | (12) | 4 ft. 4 in. = in. | | | | | | | | | | | | | | | |
|-------|--|------|---|-------|-------|------|----|---|---|-----|----|-----|----|---|---|------|----|---|
| (2) | $\begin{array}{r} 760384 \\ - 276435 \\ \hline \end{array}$ | (13) | 150 per cent of 250 = | | | | | | | | | | | | | | | |
| (3) | $\begin{array}{r} 4832 \\ \times 578 \\ \hline \end{array}$ | (14) | 14 mds. 4 seers 8 chh.
divided by 4 = | | | | | | | | | | | | | | | |
| (4) | 88)43890 (| (15) | Principle Rs. 900
Time 1 year 4 months
Rate 6 per cent
Interest = | | | | | | | | | | | | | | | |
| (5) | $\begin{array}{r} 8\frac{1}{2} \\ - 6\frac{1}{2} \\ \hline \end{array}$ | (16) | 33 $\frac{1}{3}$ per cent of 96 = | | | | | | | | | | | | | | | |
| (6) | 65 ÷ 65 = | (17) | 5 per cent of Rs. 45 = | | | | | | | | | | | | | | | |
| (7) | 2/3 + 3/4 + 5/6 = | (18) | <table border="0"> <thead> <tr> <th>Md.</th> <th>Seers</th> <th>Chh.</th> </tr> </thead> <tbody> <tr> <td>23</td> <td>4</td> <td>7</td> </tr> <tr> <td>8</td> <td>11</td> <td>4</td> </tr> <tr> <td>10</td> <td>0</td> <td>0</td> </tr> <tr> <td>+ 14</td> <td>15</td> <td>8</td> </tr> </tbody> </table> | Md. | Seers | Chh. | 23 | 4 | 7 | 8 | 11 | 4 | 10 | 0 | 0 | + 14 | 15 | 8 |
| Md. | Seers | Chh. | | | | | | | | | | | | | | | | |
| 23 | 4 | 7 | | | | | | | | | | | | | | | | |
| 8 | 11 | 4 | | | | | | | | | | | | | | | | |
| 10 | 0 | 0 | | | | | | | | | | | | | | | | |
| + 14 | 15 | 8 | | | | | | | | | | | | | | | | |
| (8) | 9 × 4 2/3 = | (19) | <table border="0"> <thead> <tr> <th>Miles</th> <th>Fur.</th> <th>Yd.</th> </tr> </thead> <tbody> <tr> <td>12</td> <td>6</td> <td>0</td> </tr> <tr> <td>- 6</td> <td>2</td> <td>110</td> </tr> </tbody> </table> | Miles | Fur. | Yd. | 12 | 6 | 0 | - 6 | 2 | 110 | | | | | | |
| Miles | Fur. | Yd. | | | | | | | | | | | | | | | | |
| 12 | 6 | 0 | | | | | | | | | | | | | | | | |
| - 6 | 2 | 110 | | | | | | | | | | | | | | | | |
| (9) | 13/5 ÷ 1 3/5 = | (20) | Give the average of 10,
6, 8, 5. | | | | | | | | | | | | | | | |
| (10) | $\begin{array}{r} 421 \\ 4.566 \\ .008 \\ .9 \\ 315.045 \\ + .17 \\ \hline \end{array}$ | (21) | The time is now 42
min. after 6-0.
What will it be in
8 hr. 30 min. from
now? | | | | | | | | | | | | | | | |
| (11) | $\begin{array}{r} 4.05 \\ \times 3.99 \\ \hline \end{array}$ | | | | | | | | | | | | | | | | | |

- | | |
|---|---|
| <p>(22) At the rate of 10 pice each, what is the price of 96 knives?</p> <p>(23) $625 \times 1000 =$</p> <p>(24) $4.25 + .05 =$</p> <p>(25) $16:48 :: 4 :$</p> <p>(26) It takes $2\frac{1}{2}$ yds. of cloth to make a shirt. How many shirts can one</p> | <p>make of a piece of cloth 48 yds long?</p> <p>(27) Write in words: 3146307.</p> <p>(28) Write in figures: One crore twenty-five lakhs four thousand seven hundred and ninety-one.</p> |
|---|---|

Each right, 1 point

Score

ARITHMETIC PROBLEMS

Time 20 minutes

Write the answers in the block.

1. If you get 20 mangoes to the rupee, what will 2 dozen cost? ()
2. Mohan went to the bazaar to get 8 seers of nails. There were 4 chhataks of nails in each packet. How many packets should he buy? ()
3. Mohan paid as. $\frac{1}{3}$ for cabbage seeds. He grew 58 cabbages and sold them for 1 anna each. How much did he earn? ()
4. The area of a square is 6.25 square yards. What are the dimensions of the square? ()
5. A piece of cloth was $10\frac{3}{4}$ yards long $1\frac{1}{4}$ of a yard was cut off. How much is left? ()
6. Ram earns Rs. 3/- per week. His living expenses are Rs. $2\frac{1}{4}$ - per week. How much can he save in a year? ()
7. A room is 35 feet by 20 feet. What will it cost to cement the floor at Rs. 5/- per hundred square feet? ()
8. A man wanted to build bunds around his rice field which was $43\frac{1}{2}$ yards long and 30 yards wide. The field was approxi-

mately square. At 2 annas per hand what will it cost him to build the bund? ()

9. The average age of 5 boys is 10 years. The average of 3 girls is 12 years. What is the average age of the boys and girls together? ()

10. Budhu took a loan of Rs. 760/-. At the rate of 6% per year what interest does he owe after 2 months? ()

Each problem right counts 2 points Score

ENGLISH TEST

Directions: Read the following paragraphs carefully and be prepared to answer questions about them. In some of the questions a choice of several answers is given, only one of which is correct. Underscore the correct answer.

Time. 30 minutes

1. Perhaps the most faithful and the brightest animal that the world has ever known is the dog. For ages, the dog has been the loyal friend of man. He has worked for his master. He has saved him in times of danger. Such devotion should be returned with kindness.

- (1) Man's ever-faithful friend is a dog, cow, cat,
 horse?
(2) The dog has been tamed for years, ages, months,
 weeks?
(3) Copy one word you read in the story that means
 strong affection - - - - - -
(4) This story is about:—
 (a) How dogs work, (b) Man's loyal friend, (c) An
 Indian dog, (d) Dogs of other lands.

2. Not long ago, a man who had been unkind to Moti walked by the elephant's cage. A peepul branch came sailing through the air. The man stooped or it would have hit his head. Moti had picked up the peepul branch with his trunk. He threw the branch at the man who had been unkind to him.

- (1) What did the elephant throw? Water, cage, trunk, branch?
- (2) What kind of man was he? Good, unkind, slow, kind?
- (3) Copy one word you read in the story that means *the elephant's nose*.
- (4) This story is about:—
 (a) A pekul branch, (b) An elephant's cage, (c) A man and an elephant, (d) The elephant's trunk.

3 Among the writers of proverbs we must rank Benjamin Franklin very highly. His proverbs show rare common sense and have become famous the world over. Here are a few of his best-known ones:

(1) Never leave that till tomorrow which you can do today.

(2) God helps them that help themselves.

(3) Early to bed and early to rise makes a man healthy, wealthy and wise.

(4) Beware of little expenses; a small leak will sink a ship.

(1) Franklin made many
 (a) expenses, (b) proverbs, (c) beds, (d) healthy men.

(2) Franklin was
 (a) sensible, (b) wealthy, (c) rare, (d) early.

(3) Write one word that means *a wise saying* . . .

(4) What proverb tells us to be careful in spending our money? (Give number only.) . . .

(5) Which proverb tells us that staying up late and sleeping long in the morning is not the best thing? (Give number.)

(6) Which proverb tells us not to delay in doing that which we can do now? (Give number.) . . .

(7) Which proverb tells us that we must do that which we want done ourselves and not expect it to be done for us? (Give number.)

4. P. L. Gupta and Co.,
Calcutta.

Dear Sirs,

Please send me by goods train the following supplies:

- 6 dozen pencils at 12 annas per dozen.
- 2 quires Elephant Brand paper at Rs. 2-4-0.
- 20 slates at 2 annas each.

Please do not send tin slates but stone slates. The paint comes off the tin slates very quickly and then they are useless. The railway receipt should be sent V.P.P.

Yours sincerely,
D. SINGH

- (1) Which item will cost Rs. 2-8-0?
- (2) Of which article are 72 wanted?
- (3) Who is dissatisfied with tin slates?
- (4) How will the money for the goods be sent?

5. When the moon passes between the sun and the earth, the sun is darkened for a short time. This is called an eclipse. At first we do not notice that the moon is in front of the sun but by the time the moon has blotted out half of the sun, the light is beginning to get dim. When the moon covers most of the sun it begins to get dark, and by the time the moon gets squarely in front of the sun it is quite dark. Birds get fooled, thinking it night, and go to sleep. This part of the eclipse takes only a few minutes however. The stars are now shining even though it be noon time. Soon the moon is passing from in front of the sun and it begins to get light again. The sunshine is getting stronger every moment. It takes the moon about two hours from the time it first gets into the path of the sunshine to get entirely out of the way again.

- (1) There is an eclipse when
 - (a) birds go to sleep, (b) when the sun gets between the earth and the moon, (c) when the stars shine, (d) when the moon gets between the earth and the sun.
- (2) The time of greatest darkness takes
 - (a) a few minutes
 - (b) until noon
 - (c) two hours
 - (d) until night

(3) Who is deceived, thinking it night?

(4) What is this story about?

(a) The eclipse of the moon

(b) The eclipse of the sun

(c) Darkness

(d) Stars shining in day time

6. There are many birds in India in the cold season that are not found here in the hot season and monsoon. These birds are called migrating birds, that is, birds that regularly spend part of the year in one country and then fly away to spend the rest of the year in another country. To move to another country to live is called 'migrating'.

Among the migrating birds found in India, the ducks are perhaps the best known. But not all ducks found in India are migrators. Several ducks stay with us all the year. All the more common non-migrating ducks are ducks which at times rest on trees. Some of them even build their nests in trees. When you see a duck which sits on a tree you may be certain that it is one of those which never leaves us all the year round. The migrators are not tree ducks, neither do they hatch out their young in India.

Many birds besides the ducks are migrators, namely cranes, some snipe, a species of stork and many song-birds.

(1) Migrators are:

(a) ducks, (b) animals or birds that travel to another country to live, (c) birds that live in trees, (d) birds that build nests

(2) Ducks that have their young in India are:

(a) migrators, (b) non-migrators, (c) common ducks, (d) ducks that never sit in trees.

(3) Name one bird that is mentioned as a migrating bird that is not a duck.

(4) This story is about:

(a) ducks, (b) cranes, (c) migrating birds, (d) tree ducks.

Each right, 1 point

Score:.....

HINDUSTANI VOCABULARY TEST

Time- 5 minutes

The meaning of the word		
2	MERH	is Mardan, Khet, Bandh, Bagicha, Kinara.
"	JIMMA	" Jawabdari, Dhan, Jabran, Imandari, Bhar.
"	LAKSHAN	" Lakshman, Lata, Chih, Kisan, Sanket.
"	SAWDHAN	" Nazar, Udyogi, Widwan, Jagrit, Sachet.
"	HASTAKSHAR	" Dastkhat, Likhawat, Akshar, Uchcharan, Lupi.
"	BAHISHKAR	" Nikalna, Dharna, Larhna, Rokna, Chhorna.
"	BHASHAN	" Kathan, Kahani, Vyakhyan, Ultha, Kawita.
"	THATH BAT	" Ghemand, Sajawat, Rasta, Bhasha, Barhappan.
"	PALA	" Tushar, Bhaf, Patthar, Garni, Os.
"	RATHOUR	" Kamzor, Shaktishali, Ek Rajwansh, Kshatri, Rajput.
"	UPARJAN	" Kamana, Jorhna, Upay, Dhandha, Mihnat.
"	SWARGBAS	" Rog, Devlok, Nark, Mrityu, Apar Sukh.
"	DUSTAR	" Saral, Asambhav, Kathin, Kathorta, Kanun.
"	TIRASKAR	" Ninda, Adar, Namaskar, Ghrina, Niradar.
"	PRANHALI	" Niti, Niyam, Riti, Pran, Dabav.
"	TATKAL	" Dhupkal, Turant, Vilamb, Kal, Deri.
"	AZADI	" Swatantrta, Gulami, Naukri, Niradhar, Abadi.
"	BHASHANTAR	" Pustak, Anuvad, Lekh, Madhy, Vaky.
"	JALAD	" Nirjal, Bijli, Badal, Adnhi.
"	CHHATR	" Manushy, Balak, Vidhyarthi, Khilarchi, Larhka.
"		"

The meaning of the word		ADHYAPAK	is	Adbyan, Skul, Headmaster, Shikshak, Parhnewala
"	"	PRATISHTHA	"	Uttam, Samman, Lobh, Dhamki, Nam.
"	"	WIHANG	"	Hans, Kauwa, Tota, Pakshi, Mor.
"	"	CHUNGI	"	Birhi, Churhi, Mahsul, Churan, Dak.
"	"	AWANI	"	Prithwi, Swarg, Patal, Parvat, Zamin.
"	"	VIPIN	"	Vayu, Samudr, Talab, Jangal, Pinna.
"	"	KALAH	"	Mitrata, Kalikal, Jhagrha, Dushman, Kal.
"	"	PIPASA	"	Bhukh, Asha, Pyas, Pyar, Lalchi.
"	"	SOPAN	"	Madira, Sharbat, Pan-supari, Charhav, Sirhi.
"	"	SAMIR	"	Hawa, Swarth, Saman, Amir, Dayalu.
Each right, $\frac{1}{2}$ point				Score

SCIENCE AND HYGIENE

Time: 20 minutes

In the second column the known effects of heat upon certain chemicals are listed. The first column is the name of the chemical. Write the number of the chemical in the parenthesis which you think is affected by heat as noted in column two.

I. Name of chemicals	Effect when heated
1. Salt	() Turns yellow when heated.
2. Paraffin	() Crystals break and a crackling sound is heard.
3. Blue Vitriol	() Not affected by ordinary heating temperature.
4. Sulphur	() Sublimates and a choking smell is produced.
5. Magnesium	() Burns with bluish-yellow flame.
6. Ammonium Nitrate	() Melts and then burns with smoky flame.
7. Chalk	() Sublimates when heated. Noticeable odour.
8. Green Vitriol	() Burns with thick smoke of bright light.
	() Burns white when heated.

II. Match the undermentioned 'Apparatus' with its use:—

Apparatus	Use
1. Callipers	() For measuring atmospheric pressure.
2. Mariner's Compass	() Chemicals are heated and tested in it.
3. Barometer	() Used by ships at sea to determine directions.
4. Graduated cylinder	() Used for measuring mass and weight.
5. Rain gauge	() Used for measuring heat.

<i>Apparatus</i>	<i>Use</i>
6. Thermometer	() Ordinarily used for measuring amounts of liquid.
7. Balance	() Used for measuring rainfall.
8. Test tube	() Used for measuring diameter.

III. Match the 'Effect' with its Cause. —

<i>Cause</i>	<i>Effect</i>
1. Bending of the sun's rays by rain drops.	() Eclipse.
2. Earth's turning upon its axis in its revolution about the sun.	() Evaporation of water.
3. Sun's rays shining hot and the air being dry.	() Floating of substances.
4. The passing of the moon between the earth and the sun	() Appearance of the rainbow.
5. Excess of water-vapour in atmosphere	() Succession of day and night.
6. Change of the position of the moon with relation to the sun and the earth.	() Appearance of the 'phases of the moon'.
7. Earth's wobbling on its axis at an angle of $23\frac{1}{2}^{\circ}$.	() Rain, snow or hail.
8. Substances being lighter than the weight of equal volume of water.	() Succession of the seasons.

IV. Match the 'Substance' with its properties:—

<i>Substances</i>	<i>Special properties</i>
1. Chalk Gas (Carbonic Acid Gas)	() Helps substances to burn
2. Nitrogen	() May readily be turned into solid, liquid and gaseous forms.
3. Oxygen	() When introduced into colourless lime-water, it turns it milky.
4. Sawdust	() Gives off oxygen in presence of sun's light.
5. Phosphorus	() Even though it is a liquid, a solid iron-ball does not sink in it.
6. Vegetation	() Bad conductor of heat.
7. Mercury	() In its presence ice does not melt at the usual temperature
8. Water	() Burns when it comes in contact with air.

V. Underline the best answer to the statements given below:—

- Only those clothes should be considered clean
 - which get white when washed.
 - which are washed with one's own hands.
 - which are boiled and washed with soap.
 - which are washed in river water.
- It is desirable to bathe daily
 - because it is a religious rite.
 - to keep the body cool.
 - to keep the pores of the skin open in order to allow the inside dirt to pass out with perspiration.
 - for appearance sake.

*Substances**Special properties*

- | | |
|--|---|
| <p>3. We should not let any one else enter the room of a smallpox patient other than the one attending him or her</p> <p>4. The most necessary part of the treatment for a cut of the skin is to</p> <p>5. A house may be described as properly cleaned when</p> <p>6. It is necessary to clean the finger nails before taking meals</p> <p>7. In order to stop typhoid from spreading</p> | <p>(a) because it might make the patient angry.</p> <p>(b) because smallpox is an infectious disease.</p> <p>(c) because the entrants might be sorcerers.</p> <p>(d) because the patient needs rest.</p> <p>(a) tie it up with a piece of cloth.</p> <p>(b) wash it with disinfected water.</p> <p>(c) paint it with tincture iodine.</p> <p>(d) smear dust over it.</p> <p>(e) apply lime.</p> <p>(a) the sweepings have been thrown out.</p> <p>(b) the sweepings are piled at one place.</p> <p>(c) the surroundings are also thoroughly cleaned</p> <p>(d) the house has been plastered with cow-dung.</p> <p>(e) the house has been swept many times a day</p> <p>(a) because they may be dirty.</p> <p>(b) because there is a sort of poison in the nails.</p> <p>(c) because there is a sort of liquid produced in nails.</p> <p>(d) because it is a good custom.</p> <p>(a) patients should be given clove water.</p> <p>(b) carefully dispose of the patient's stool and urine.</p> <p>(c) patients should be given proper medicine.</p> <p>(d) inoculate the patient.</p> |
|--|---|

8. What organ rids the body of poison?
- (a) Kidneys.
 - (b) Lungs.
 - (c) Heart.
 - (d) Brain.
 - (e) Liver.
-

Each right, $\frac{1}{2}$ point

Score

.

MIDDLE SCHOOL ACHIEVEMENT TEST
NORMS

	Hand reading 23*	Arith. comp. 28*	Arith. prob. 20*	Engl. 27*	Total on Part I only 98*	Vocab. 15*	Grammar usage 20*	History 25*	Geo- graphy 20*	Science 20*	Total 198*
Class VII											
Q3	18.5	17.8	10.1	12.5	58.9	9.8	11.9	11.9	9.9	10.5	101.9
Median (Norm)...	16.1	14.3	8	9.2	47.6	8.4	9.5	10.3	7.6	8.8	92.2
Q1	13.3	10.6	4.5	5.2	33.6	6.4	7	8.6	6	6.5	68.1
Class VIII											
Q3	20.1	20.1	12.8	18.5	71.5	11.9	15.2	15.3	11.7	13.5	139.1
Median	17.7	17	10.1	15	39.8	10.5	12.8	13	9.7	11.2	117.8
Q1	14.8	13.3	6.4	9.5	44.0	8.5	10.5	11	7.7	8	89.7
Class IX											
Q3	21.5	20.4	12.6	22.7	77.2	12.5	15.5	15	11.5	15.1	146.8
Median	19.4	17	10	19.8	66.2	11.1	13.5	12.6	9.6	13.6	126.6
Q1	17.5	13.5	6.5	15.5	55.0	9.5	11.4	10.5	8.4	11	103.8

*Indicates maximum number of points possible.

APPENDIX III

THE CHATTERJI ARITHMETIC TEST

This test is reprinted with the kind permission of the author. It is intended for high schools.

I. MECHANICAL ARITHMETIC

Time allowed— $\frac{1}{2}$ hour.

Attempt only those sums that you know. Put down your answers at the end of the questions.

Work mentally as far as you can; scrap work is, however, allowed on separate paper.

1. Express seven lakhs thirteen thousand and two in figures.
2. $9+6+12=$
3. $769-635=$
4. Express 190805 in words.
5. $1 \div 1=$
6. Turn 2097 pies into Rs. as. ps.
7. Subtract Rs. 2-5-3 from Rs. 17.
8. $19-5-3=$
9. $2000-1=$
10. Turn $\frac{1}{2}$ into a decimal.
11. $(.09)^2=$
12. What is x in $4\frac{x}{7}=1\frac{4}{7}$?
13. $\sqrt{576}=$
14. $444-400=$
15. $57 \times 8=$
16. $6 \times 7 \times 0 \times 3=$
17. Add:

Rs.	as.	ps.
17	14	3
6	15	9
8	8	9
18. What is x in $\frac{3}{7} = \frac{21}{x}$?

19. Which one is greater, $\frac{3}{4}$ or $\frac{4}{5}$?
20. $4\frac{1}{2} + 3\frac{1}{2} - 2\frac{1}{2} =$
21. $6\frac{3}{8} = 2\frac{7}{8} =$
22. $.5 + .75 + .25 =$
23. $4 \div .004 =$
24. Rs. $13 - 6\frac{1}{2}$ as. $\times 2 =$
25. Find the compound interest on Rs. 150 for 2 years at the rate of 4% per annum.
26. Find the simple interest on Rs 250 for 2 years at 5% per annum.
27. $2942 \div 6 =$
28. $0 \div 4 =$
29. $72 \div 72 =$
30. Turn 7 mds. 30 or 15 chh. into chhataks.
31. 15 yds. 2 ft. 6 in. $- 5 =$
32. $8 \times 2 - 21 \div 7 =$
33. Rs. $17 - 4$ as. \div Re. $1 - 7$ as. $=$
34. Find the G.C.M. of 16, 24, 56.
35. Find the L.C.M. of 1, 2, 3, 4, 5.
36. Reduce $2\frac{1}{3}$ to its lowest terms.
37. Find the number of days in the first three months of the year 1932.
38. $\frac{4}{7} \div \frac{1}{21} =$
39. Simplify $\frac{1}{4\frac{1}{2} \times 2}$ of £ 22.
40. $1 \div 2\frac{1}{2} =$
41. $\frac{1}{2} \div \frac{1}{4} \times \frac{1}{2} =$
42. $\frac{7 \text{ days } 4 \text{ hrs.}}{1 \text{ day } 19 \text{ hrs.}} =$
43. Express .0001 as a fraction.
44. Express $1\frac{25}{100}$ as a decimal.
45. $5 - 3.3 =$
46. What is 5% of 10 shillings?
47. What is 1% of 100?
48. What is 100% of 100?
49. What is 100% of 1?
50. Find the average of 5, 6, 7, 10.

II. ARITHMETICAL REASONING

Attempt only those sums that you know. Put down your answers at the end of the questions.

Work mentally as far as you can; scrap work is, however, allowed on separate paper.

1. Two thirds of a class is 20 boys. How many are there in the class?
2. Find the cost of 12 chairs at Rs. $5\frac{1}{8}$ per pair.
3. A carpet 20 ft by 15 ft. costs £ 10. What was the cost per square yard?
4. Ram is as old as Mohan; Mohan is twice as old as Rahim; Rahim is as old as Hari. If Hari is 5 years old, how old is Ram?
5. I want equal numbers of stamps and postcards for Re. $1\frac{1}{5}$. If a stamp costs one anna and a postcard 9 pies, what will be the number of each?
6. How many men will do a piece of work in 12 hrs. that 4 men can do in 18 hrs.?
7. A man takes 20 minutes to walk from his house to the station. His son also takes 20 minutes. How long will it take them if they both walk together?
8. What is the greatest number that will divide 64 and 76, and leave a remainder 4 in each case?
9. The area of the floor of a room is 156 square feet. What is the sum of its four sides?
10. If mangoes sell at the rate of Rs. $3\frac{1}{2}$ per 100, how many can I buy for 7 as.?
11. The distance round the wheel of a motor car is 3 ft. 6 in. How many times does the wheel go round in travelling 140 yds.?
12. Find two numbers whose sum is 19 and whose difference is 5.
13. The area of a field twice as long as it is broad is 200 sq. yds. Find its sides.
14. How many times can I subtract 3 from 73?
15. What is the length of a stick which I can cut up into 8 pieces, each 6 in. long and have 4 in. left over?

16. A man aged 35 years is 7 times as old as his son; how many times as old as his son will he be 25 years hence?

17. At what rate per cent simple interest, would Rs. 10 amount to Rs. 11 in 4 years?

18. Rs. $\frac{9}{11}$ is made up of equal numbers of rupees, eight-anna, four-anna, two-anna and one-anna coins. Find the number of each.

19. I buy some articles for Rs. 40; what must I sell them for, so as to gain 20%?

20. If $\frac{3}{5}$ of a piece of cloth cost Rs $\frac{3}{6}$, what is the cost of the whole piece?

21. After spending half of my money and then half of the remainder I had 2 annas left. How much had I at first?

22. Divide Rs. 30 between A and B so that their shares may be proportional to 6 and 4.

23. What is the least number which must be subtracted from 45 to make it exactly divisible by 7?

24. What two whole numbers multiplied together make 11?

25. There are two numbers one of which is greater than the other by 4. When multiplied together they make 165. What is the smaller one?

26. If telegraph poles stand 50 yards apart in a straight row, what is the distance from the first to the eighth?

27. 5 per cent of A's income is the same as 15 per cent of B's. A's income is Rs. 300 a year. What is B's?

28. A brick weighs 7 lb and half its own weight. What is the weight of the brick?

29. If a man's salary is Rs. 80 a month and he spends Rs. 56 a month, how long will it take him to save Rs. 600?

30. If 2 pencils cost 5 pice, how many pencils can you buy for 50 pice?

Notes: — Printed test papers both of the mechanical and reasoning arithmetic, answer-key, and instructions will be sent on application at a nominal price. The applicants are requested to mention the Indian language in which they want the above materials.

Patna Training College.

NORMS

	<i>Mechanical Arithmetic</i>	<i>Arithmetical Reasoning</i>
Class VIII	22	10
Class IX	25	11
Class X	29	13.6
Class XI	34	16.6
Graduate Adults		23

APPENDIX IV

A 'CHARACTER CUP' FOR CLASS VIII

by M. P. DAVIS, D.D.

ALL boys give their estimate for every boy in the class. It is assumed that the boys (or girls) have been together in school for several years and know each other well. Their estimates are strictly confidential and they should be advised not to tell any one else just what they entered. It should be explained to them carefully that the estimate is to be made very objectively, that is, the estimates should be made strictly on the merits of each boy. No boy should be given a good mark just because he is a special friend or be given a bad mark because he is not liked. Every voter must be as fair as he can possibly be and judge each boy on his own merits on every item separately. No marks at all should be given in cases where the voter is too uncertain to vote intelligently.

If the boys of a certain class are well known to the boys of other classes, others may also give their estimates but it is to be remembered that in so far as the ones to be voted on are *not well known* to each other the Character Cup loses its reliability. As the boys in the highest class are nearly always the best known in the school, they can be voted on by other classes in schools where the enrolment is not over 150. There should be at least 25 people voting on each other.

Copies may be cyclostyled with the names of all boys on which estimate is to be made included in the list. The scoring is done by simply adding the number of marks given by all those voting and dividing by the number of boys who voted for him. A separate estimate by members of the staff for comparison with that of the boys will prove interesting. Boys usually know more about each other in such matters than the teacher is apt to know about them. A naturally popular boy in the school is apt

APPENDIX V

SINCE the system of naming classes is not consistent in the various provinces and states of India, it is necessary to explain the system used in this book. The Central Provinces system is used. This is the simplest. Classes I to IV comprise the primary school, classes V to VIII the middle school and classes IX to XI the high school. We might use the term 'year' just as easily: Class I is Year I, Class V is Year V, and so on. In other provinces Kindergarten or Infant Classes precede the regular classes, standards or forms. In Madras there are both standards and forms, the latter signifying entry into the middle school. Class III in the C.P. and Punjab are by no means equivalent, for Class I is the first year of school in the C.P. while it is the 3rd in the Punjab. In the following table of class systems it will be seen that there is not only difference in terms and numbering but that the number of classes in the primary school system, middle school system, and high school system varies greatly. It is a very confusing situation.

To repeat, the system used here is the Central Provinces system in which Class I is the first year of school, Class IV the fourth, etc. When making comparisons with the school system of your own province or state, or when considering standards and norms for various classes, the equivalent class placements for your own province have to be worked out.

INDIAN CLASS SYSTEMS, I. I. 41.

	PRIMARY	MIDDLE	HIGH
Assam *	Cl. I-IV	Cl. V-VI	Cl. VII-X
Baroda	In., St. I-IV	St. V-VII	St. IV-VII
Bengal	Cl. I-IV	St. I-III	Cl. VII-X
" Eur	In., Cl. I-IV	Cl. V-VI	Cl. VII-X
Bihar & N. Or.	Cl. I-IV	Cl. VI-VII	Cl. VIII-XI
Bombay	In., St. I-IV	St. V-VII	St. IV-VII
" Eur.	Kn. I-II, St. I-IV	St. I-III	St. V-VII (J. C.)
Burma A. V.	St. I-IV	St. V-VII	St. VIII-IX (S. C.)
" C. N. E.	Kn. Fm. I-VII		St. VIII-X
" Eng.	In., St. I-IV	St. V-VII	Fm. VIII-XI
Central Prov.	Cl. I-IV	Cl. V-VIII	Cl. IX-XI
Ceylon	Kn. St. I-V	St. VI-IX	St. X-XI
Cochin	Cl. I-IV	Fm. I-III	Fm. IV-VI
Hyderabad Osm.	Cl. I-IV	Cl. V-VIII	Cl. IX-X
" H. S. L. C.			Fm. IV-VI
Kashmir	In., Cl. I-V	Cl. VI-VIII	Cl. IX-X
Madras	St. I-V	St. VI-VIII	Fm. IV-VI
		Fm. I-III	

INDIAN CLASS SYSTEMS, 1 : 41 — (continued)

	PRIMARY	MIDDLE	HIGH
Mysore	St. I-VI	St. VII-IX	Fm. IV-VI
Punjab	In., I-II, Cl. I-IV	Cl. V-VIII	Cl. IX-X
Rajputana	In., I-II, Cl. I-IV	Cl. V-VIII	Cl. IX-X
Sind	In., St. I-IV	St. V-VII	
S. Orissa	St. I-V	St. I-III	St. IV-VII
Travancore	Cl. I-IV	Fm. I-III	Fm. IV-VI
United Provinces	In., Cl. I-IV	Prep. Fm. I-III	Fm. IV-VI
W. I. States	In., St. I-IV	Cl. V-VIII	Cl. IX-X
		St. V-VII	
		St. I-III	St. IV-VII

APPENDIX VI

BIBLIOGRAPHY

THIS list is purposely confined to works easily obtainable and usable in India.

ON TESTING AND MEASUREMENTS (FOREIGN)

- * Ballard, P. B. *Group Tests of Intelligence*. (University of London Press,—)
- * Ballard, P. B. *Mental Tests*. (University of London Press,—)
- ‡ Ballard, P. B. *The New Examiner*. (University of London Press, 1936.)
- ./ Bingham, Walter Van Dyke, *Aptitudes and Aptitude Testing*. (Harper and Brothers, 1936.)
- * Burt, Cyril. *Handbook of Tests*. (P. S. King, London, 1923.)
- ** Burt, Cyril. *Mental and Scholastic Tests*. (London City Council, 1921)
- ✓ Cattell, R. B. *A Guide to Mental Testing* (University of London Press, 1936.)
- Freeman, F. N. *Mental Tests*. Revised edition. (Harrap and Co., London,—)
- Garrett, H. E. *Great Experiments in Psychology*. (The Century Co., New York, 1930.)
- Goodenough, F. L. *Measurement of Intelligence by Drawings*. (Harrap and Co., London, 1926.)
- Hamley, H. R. *The Testing of Intelligence*. (Evans Bros., London, 1935)
- ✓ Hawkes, H. E. *Achievement Examinations (The Construction and Use of)*. (George G Harrap and Co., 1937.)

* Those marked with an asterisk have a good many test examples which will prove suggestive. ** Those marked with two asterisks are the ones the author has found most suggestive.

- Hull, C. L. *Measurement and Adjustment Series*, (World Book Company, New York, 1928.)
- Hunt, E. P. A. and Smith, P. *The Teachers Guide to Intelligence and Other Psychological Testing*. (Evans Bros., London, 1935.)
- Kandel, I. *Examinations and their Substitutes*. (Carnegie Foundation for the Advancement of Learning, New York, 1936)
- Knight, Rex *Intelligence and Intelligence Tests*. (Methuen and Co., London, 1933)
- McCall, Wm. A. *Measurement*. (The Macmillan Company, New York, 1939.)
- Pintner, R. *Intelligence Testing*. (Henry Holt and Co., New York, 1930.)
- *† Rinsland, H. Daniel. *Constructing Tests and Grading*. (George G. Harrap and Co., London, 1938.)
- Stutsman, R. *Mental Measurement of Pre-School Children*. (World Book Co., Chicago, 1931.)
- Symonds, P. M. *Diagnosing Personality and Conduct*. (Century Co., New York, 1930.)
- Symonds, P. M. *Measurement in Secondary Education*. (Macmillan, New York, 1930.)
- Technical Staff of the Board of Examiners of the University of Chicago. *Manual of Examination Methods* (University Press, Chicago, 1933.)
- Terman, L. *The Measurement of Intelligence*. (Houghton Mifflin, Boston, 1916.)
- Terman, L. A later edition of the above is entitled: *Measuring Intelligence*. (Harvey, 1937.)
- Thorndike E. L. *The Measurement of Intelligence*. (Columbia University, New York, 1926.)
- Wiggam, A. *Exploring Your Mind*. (Bobbs Merrill Co., Indianapolis, U.S.A., 1938.)
- Yoakam, C. S., Yerkes, R. M. *Army Mental Tests*. (Henry Holt and Co., New York, 1920.)

STATISTICS

- Brown, W. and Thomson, G. H. *Essentials of Mental Measurements* (Cambridge University Press, 3rd Ed., 1925.)

- Rugg, H. *A Primer of Graphics and Statistics for Teachers*. (Houghton Mifflin, New York, 1925.)
 Rugg, H. *Statistical Methods Applied to Education*. (Houghton Mifflin, New York, 1917.)
 Thurstone, L. L. *The Fundamentals of Statistics*. (Macmillan, New York, 1925.)
 Tieg, E. W., and Crawford, C. C. *Statistics for Teachers*. (Houghton Mifflin, New York, 1930.)

BIBLIOGRAPHY OF INDIAN PUBLICATIONS

- Ahmed, Z. U. *Systems of Examinations*. (Longmans, Green and Co., 1932.)
 D'Sylva, L., and stat. *The Measurement of Attainments in Primary Schools*. (Government Printing, Nagpur, 1937.)
 Jha, B. N. *Modern Educational Psychology*. (The Indian Press, 1933.)
 Kamat, V. V. *Measuring Intelligence of Indian Children*. (Oxford University Press, 1940.) In Marathi and Kanarese also.
 Kumria, R. R. *Intelligence, Its Nature and Measurement*. (Gulab Singh and Sons, Lahore, 1936.)
 Mathew, A. K. *Examinations, A Constructive Survey*. (Gulab Singh and Sons, Lahore, 1934.)
 Menzel, E. W. *Suggestions for the Improvement of Reading in Indian Secondary Schools*. (Oxford University Press, 1942.)
 Perera, H. S. *The Psychology of Learning and Teaching*. (Oxford University Press, 1937.)
 Rice, C. H. *A Hindustani Binet-Performance Point Scale*. (Oxford University Press, 1929.)
 Ryburn, W. M. *The Progressive School*. (Oxford University Press, 1938.)
 Teachers' College, Saidapet. *Bulletin No. 15*. (Government Press, Madras, 1923.)
 Woodburne, A. S. *Psychological Tests of Mental Abilities*. (Government Press, Madras, 1924.)

(The last named work is an able and comprehensive presentation meriting much more attention than it has received. It predates all but one or two small experiments in India.

INDEX

- Accuracy, standards in arithmetic, 62.
- Achievement Test in Middle School subjects, 66, Appendix II
- Achievement tests (*see also* school-subject tests), 43, 66.
- Adaptation of intelligence tests, 154.
- Advantages of New-Type tests, listed and discussed, 88-102.
- Analysis of pupils' difficulties (*see* Diagnostic tests, Chapter V), 111.
- Aptitude tests, Chapter VIII, 180.
 aptitude tests defined, 182.
 purpose of, 182.
Stenquist Assembly Test, 183.
MacQuarrie Mechanical Aptitude Test, 184.
 still experimental, 228.
- Arithmetic Tests, 59-68.
 Two types of arithmetic to be tested, 60-62.
 speed tests and power tests, 61-62.
Northumberland tests, 63.
Monroe Survey Tests, 64.
 The author's, 65.
- Chakravarty Scale*, 67-68.
- Army Intelligence Tests, (*Army Alpha* and *Army Beta*) 141-42.
- Ayer's *Handwriting Scale*, 58.
- Ballard, P. B.
 One-Minute Reading Scale, 46-48.
 Silent Reading Test, 52.
 where to get tests, 83-84.
 appraisal of intelligence tests, 146.
 intelligence tests, 151.
- Bhatapara Middle School, Appendix II.
- Binet, A.
 experiment in measuring intelligence, 133-39
 analysis of intelligence, 143.
 measuring the intangibles, 191.
- Binet - Simon Intelligence Test, 54.
 history of, 136-39.
 effects of environment on test, 157.
 Urdu and P u n j a b i version, 159.
 Marathi and Kanarese version, 161.

- considered most valid, 162.
 disadvantages, 162.
 trustworthiness, 178.
 Brown, W., 143.
 Burt, C., 171.
 Northumberland Arithmetic Test, 63.
 Mental and Scholastic Tests, 83.
 version of Binet-Simon test, 139.
 Buswell, G. T., diagnostic chart, 117.

 Calcutta University, 162, 171.
 Cattell, J. M., 134.
 Cattell, R. B., 151.
 Chakravarty, S. C., Problems Scale, 66, 67.
 Chance element in tests, 91-93.
 Character Inventory, 192, 196.
 Chatterjee, N. C., 59, 66.
 Chichermane, D. V., 65.
 Chilson, E., 176.
 Chronological age, 138.
 Classification test, 19-21.
 Clerical Ability Test, 185.
 Coaching in tests, 106.
 Coefficient of correlation, 220.
Compass Diagnostic Test, 118.
 Completion Tests, 17.
 directions, 18.
 chances in guessing, 19.
 Computation tests, 60-66.
- diagnosis in comp.*, 115-18.
 Conduct tests, 193.
 Consistency of a test, (*see also* reliability) 39.
 Correlations, 220.
 Cost of standardized tests, 108, 224.
 Cox, J. W., 185.
 Crawford, C. C., Statistics, 205.

 Dacca Training College, 162.
 Davis, M. P., Appendix IV.
 Diagnosis, in old-and new-type tests compared, 104-105.
 in computation, 115-18.
 in problems, 118-20.
 in reading, 120-22.
 Diagnostic tests, use of, Chapter V, 111.
 purpose of, 112-15.
 to locate specific difficulties, 112-17.
 Buswell-Johnson Test, 118-19.
 Compass Tests, 118.
 Stevenson Test, 118.
 Pressey Reading Tests, 121.
 differentiated from instructional tests, 122.
 evaluated, 233.
 Dictation, poor method of teaching spelling, 113.
 Dougherty, B. L., 196.
 Draper, M., 176.
Draw-a-Man Test, 171.
 D'Sylva, L., 171, 232.

- Errors, in multiplication, 117-18.
- Essay-type of examination, 89.
- scoring subjective, 89.
- limits speed of answering, 96.
- not obsolete, 102.
- function of, 102.
- Ewing, R., 176
- Examination of Examinations*, 36, 89.
- Expressional work, 103.
- Folios of tests, 86.
- Formal discipline, doctrine of, 144.
- Frequency distribution, 207-8.
- Frequency distribution sheets, 207-8.
- Galton, G., 134-43.
- Garrett, H. E., 144.
- General Intelligence Test for Africans*, 175-77.
- Goodenough Intelligence Test*, 32, 171-74, 178.
- illustrations, 173.
- Gray Oral Reading Tests*, 45.
- Gray Oral Reading Check List*, 45.
- Grading of curriculum (influence of New-Type Tests), 100-2.
- Group Test of Mental Ability*, 170.
- Group tests in intelligence, 163.
- need of, 178.
- Guessing,
- in True-False tests, 7.
- in Multiple Choice tests, 14.
- in Matching tests, 15.
- in completion tests, 19.
- 'Guide to Mental Testing', 151.
- Hamley, H. R., 120.
- Handbook for Teachers*, 76.
- Hartog, P., 35.
- Hartshorne, H., 193.
- Henmon Latin Test*, 78.
- Hillegas Composition Scale*, 73.
- How to construct True-False tests, 3.
- Hudelson Composition Scale*, 73.
- Hunt, E. P. A., 195.
- 'Improvement of Reading Habits', 48, 52, 120.
- Individualized instruction, 123.
- necessary in diagnostic work, 123.
- Inequality of tests, 97.
- Informal Classroom tests, Chapter I.
- usefulness, 230.
- Intelligence,
- importance of, 130.
- distribution of, 132.
- measurement difficulty, 133.
- use of word 'Intelligence', 140.

- nature of, 143-45
 indication of, 156
 Intelligence Quotient, 139
 on Hindustani Binet
 Scale, 161.
 should pupil be told his
 IQ? 229.
 Intelligence tests, Chapters
 V and VI, 129
 purpose of, 130-33.
 history of, 133-43.
 Binet-Simon Test, 135-
 40.
 Army Intelligence Tests,
 141.
 uses of intelligence tests,
 147-50.
 General Intelligence Test
 for Africans, 147, 175-
 76.
 list of English and
 American tests, 151-
 53.
 use in India, Chapter
 VII, 154.
 inadequate for race com-
 parisons, 157.
 Binet Tests, 159, 162.
 lists of Intelligence tests
 in India, 163-77.
 Instructional tests, Chapter
 V, 111.
 differentiated from diag-
 nostic, 122.
 Inventories, 196-99.
 Character Inventory, 196.
 Psychoneurotic Inven-
 tory, 198.
 IQ (*see* Intelligence Quot-
 ient)
- Jalal Udin, 170.
 Jalota, S., *Intelligence*
 Test, 165-70
 Jha, L. S., *Simplex In-*
 telligence Test, 164.
 John, L., *Diagnostic*
 Chart, 118.
 Kamat, V. V., *Marathi and*
 Kanarese Binet Scale,
 161.
 correlation with, 173.
 Karachi Training College,
 171.
 Key for evaluating papers,
 9, 94.
 Koenig, J. C., 75.
 Kumria, R. R., 170.
 Kuppaswamy Aiyangar, 90.
 Lady Willingdon Training
 College, 162, 171.
 Language Tests, 70-83.
 Teaching of mother-
 tongue and foreign
 language distinguish-
 ed, 70
 Sones-Hurry H. S. Test,
 71-73.
 Metropolitan Achieve-
 ment test, 73.
 Hillegas Composition
 Scale, 73.
 Hudelson Composition
 Scale, 73.
 Hannon Latin Test, 78.
 White Latin Test, 78.
 English Test by Leeder,
 79-81.
 An Indian Language as
 second language, 83.

- Leeder, Geo.,
 English Test, 79-81.
 Intelligence Test, 177.
Leiter International Scale,
 163.
 Length of a test, effect on
 reliability, 40, 91.
 Limitations of New-Type
 tests, 102-8.
 List of Publishers of tests,
 86-87.

*Mahalanobis Composite
 Tests*, 171.
Mait's Tests, 162, 171.
 Mannix, M. E., 196.
 Marry, J., *Preliminary
 Classification test*, 163.
 *Professional test for
 Teachers*, 188.
 Matching test, 15-17.
 directions, 15.
 chances in guessing, 15.
 May, M., 193.
Maze Test, 157.
 Mean, 209, 211.
 Measurement, need of
 better technique, 232.
 'Measurement in Second-
 ary Education', 9, 18-19.
 Advantages and limita-
 tion of New-Type
 tests listed, Chapter
 IV, 88-110.
 Measurement of Attain-
 ments, 232.
 Median, 208 f.
 Mental Age, defined, 138-
 40.
 Mental Electives, Binet's
 experiment, 133-35.
- Metropolitan Achievement
 Test*, 73.
Minnesota Aptitude Test,
 185.
 Mohsin, S. M., 170.
Mouroe Survey Test, 64.
 Moral judgement and
 knowledge, tests of, 201.
 Multiple Choice test, 11-15.
 directions, 12.
 chances in guessing, 14.

 National Institute of In-
 dustrial Psychology, 185.
 National Society for Study
 of Education, 48.
 Non-Linguistic Intelligence
 tests, 154.
 Non-verbal Intelligence
 tests, 154-55.
 examples, 171.
 Normal child, Binet's list
 of abilities, 135-38.
 Normal Probability Curve,
 215.
 illustration, 217.
 Norms, defined, 35.
 kinds of, 37.
 usefulness, 100.
*Northumberland Arithmetic
 Test*, 63.

 Objective tests, 34, 88, 95.
 Olcott, M., 177.
 Oliver, R. A. C.
 use of intelligence tests,
 147-50.
 use of Maze test, 157.
 General-Intelligence Test,
 175-77.
*One Minute Reading
 Scale*, 46-48.

- O'Reilly, F. L., 196.
Otis Self-Administering Test, 33.
 part in Army intelligence tests, 141.
- Pul's Binet Test*, 162.
Pasur Group Intelligence Test, 170.
 Passing mark on New-Type Tests, 27, 98.
 Patna Training College, 162.
 Pearson, K., 143.
 Percentile, 211.
 Percentile system of determining 'passes', 99.
 Pereira, H. S., 170.
 Performance tests, difficulties in character testing, 193, 204.
 Personality tests, list of, 192.
 Phonetics, in Indian languages, 113.
 Pinnei, R., 151-78.
 Power Tests, 61, 64.
 Probable Error, 226.
Preliminary Classification Test, 163.
Pressey Diagnostic Tests, 121.
Pressey XO Tests of Emotions, 197.
 Primary School Test, Appendix I.
 Problems, Arithmetic, 60-61.
 Professional Test for teachers, Manry, 188.
 Shah, 189.
- Prognosis tests, Chapter VIII, 180.
 defined, 180.
 how composed, 186-87.
 professional teaching tests, 188-89.
 still experimental, 228.
- Psychological tests, 43.
 differentiated, 129.
 list of some more recent types, 192, 193.
 still experimental, 228.
- Psychoneurotic Inventory*, 198.
- Public examination, unreliability of, 89-90.
 influence on teaching, 99.
 purpose different from that of Standardized tests, 106.
- Publishers of tests, 85.
- Q, Q₁, Q₂, Q₃, 210, 211.
 Quartiles, 210.
 Questionnaires, 193, 200.
- Rahim, 81.
 Ram, Pats, 188.
 Range, 215.
 Rank Order, how to obtain, 206.
 Rating Scales, 193, 195.
 example, 195.
- Reading difficulties, diagnosed, 120.
 treatment, 125.
 in Intelligence tests, 155.
- Reading Tests, 46-56.
 Aims of teaching reading, 48.

- Weakness of traditional method of testing reading, 44-45.
- What a reading test should measure, 46-48.
- Oral Reading tests, 45.
- (*Gray's and Ballard's*)
- Silent reading tests, 47-56.
- Sangren - Woody Test*, 49-52.
- Ballard's, 51.
- The author's, 52-56.
- Diagnostic tests in reading, 120.
- Reasons for failure in problems, 120.
- Reliability, defined, 38.
- how obtained, 39
- mathematical expression, 41.
- effect of length on reliability, 41.
- Remedial instruction, 111.
- sometimes remedies very devious, 113.
- individualized instruction necessary, 123.
- Rice, C. H., Adaptation of Binet Scale, 159-61
- Richardson, C. A., *Simplex test*, 164.
- Rote-memory type of learning, 99-100.
- Rugg, H., *Primer of Statistics*, 205.
- Ryburn, W. M., 59
- Sadhu, B., Appendix I.
- Samples of School-Subject tests, where to get, 83-87
- Sangren, P. V., 49.
- Scholarship examination, suggestion for, 177.
- Scholastic tests (*see also* School-subject tests), 43.
- Schonell, F. J., 120.
- School-subject tests, Chapter III
- Where to get samples, 86.
- diagnostic features in, 126,
- usefulness, 230.
- Scoring tests, quick method, 9, 93.
- S. D. (*see* Standard Deviation)
- Seashore Musical Talents Test*, 181.
- Secrecy by examiners, 95-96.
- Shah, L. K., Intelligence test, 170.
- Shikarpore, 162.
- Short tests, apt to be unreliable, 41, 91.
- Shuttleworth, E. K., 193.
- Simon, Th., 135.
- Simplex Intelligence Test*, 164.
- Smith, P., 195.
- Sones-Harry High School Achievement Tests*, 71-73.
- Spearman, C., Part in statistical refinement, 143.

- nature of intelligence, 145, 186.
- Specific nature of pupils troubles, 114
- Speed, in reading, 50, 54
in writing, 59.
in arithmetic, 61, 62.
- Spelling, not to be taught through dictation, 113.
- Spence Training College, 171.
- Standard Deviation, 218.
- Standard Error, 226.
- Standardized Tests, Chapter II.
definition, 31.
form, 31.
grading, 32.
printed forms, 33.
scoring objective, 34.
norms necessary, 36.
development, 227.
- Statistical Devices, Chapter X, 205.
- Stenquist Aptitude Tests*, 183
- Stevenson, P. R., *Arithmetic Reading test*, 118-19.
- Survey tests, 115.
- Symonds, P., 19.
Advantages and Limitations of New-Type tests, Chapter IV.
quotations, 204.
- Teachers' and Authors' Word List*, 75-76.
- Teacher's Word Book (*see also Thorndike*), 75.
- Teaching function of tests, Chapter V, 84, 96, 100.
- Terman, L., Version of Binet-Simon Scale, 139.
popularized IQ, 139.
upward limit of mental age, 140.
- Comparison of Urdu and English Binet Scales, 161.
- Test of mental ability, 165.
- Term papers, 104.
- Thomson, G. H., 143.
- Thorndike, Ed., 75, 79.
Analysis of intelligence, 143.
- Thurstone's Scale of Attitudes toward the Movies*, 201.
- Tiegs, E. W., Statistics, 205
- Transfer of training, 144.
- Translation of tests, difficulties, 155.
- True-False tests, 3-12.
directions, 4.
pitfalls, 9.
chances in guessing, 7, 35.
- Twenty-fourth Yearbook of N.S.S.E.*, 48
- Union Christian Training College, 176.
- Unreliability of short tests, 91.
- Validity, defined, 41.
how determined, 42.
- Verbal intelligence test, 154, 155.

- Vernon, P. E., 203.
 Vicary, T. C., 176.
 Vocational Guidance tests,
 185.
 Warnshuis, J. H., 177.
 Watson, G., 142.
 Weidemann, C. C., 4.
White Latin Test, 78.
 Wiggam, A. E., 204.
 Woodworth, Transfer of
 training, 144.
 Psychoneurotic Inventory, 198.
 Woody, C., 49.
 Word counts, 74-77.
- Writing Scales, 56-59.
 what to test, 57.
 Hindi writing scale, 57.
 Ayer's Handwriting
 Scale, 58.
 how to construct, 58.
 speed standards, 59.
 Writing in tests is tiring,
 95.
 Wundt, W., 134.
 XO tests, 21.
 Yerkes, part in Army-
 Intelligence tests, 141.
 'Yes and No' tests, 11.